CHARACTERIZATION OF THE DNA-BINDING ACTIVITY OF THE *SACCHAROMYCES CEREVISIAE* TRANSCRIPTIONAL ACTIVATOR GCR1P

BY

MICHAEL ANDREW HUIE

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1994

ACKNOWLEDGMENTS

I would first like to thank my parents for their encouragement and support throughout the course of my training. I am especially indebted to my mentor, Henry V. Baker, for his patience, and for providing me with an excellent training and a stimulating work environment. The members of my committee, Daniel J. Driscoll, Richard W. Moyer, and Thomas P. Yang, are thanked for their invaluable insights. John Olson, Michael Brooks, Michael Waters, Jeff Smerage, Lucia Eisner, Carolyn Drazinic, Ed Scott, Clive Stanway, Jeff Harris, Mary-Catherine Bowman, Jim Anderson, Rob Nicholls, Gerry Shaw, Maurice Swanson, Jace Dienhart, William Hausworth, Al Lewin, Paul Gulig, Cecila Lopez, Didi Gravenstein, and Andy Wilcox are thanked for uncountable hours of discussion related to this work and for technical assistance and advice. This work is dedicated to Matt Memolo, who died during the course of the writing. He is thanked for his inspiring conversations and for his example of how to live life.

TABLE OF CONTENTS

	page
ACKNOWLEDGMENTS	ii
LIST OF TABLES	v
LIST OF FIGURES	vi
ABSTRACT	viii
INTRODUCTION	1
Basal Transcriptional Machinery and TFIID	4
Activators, Coactivators, and Adaptors	7
History of Fermentation in Yeast	11
Coordinate Regulation of Glycolysis in *Saccharomyces cerevisiae*	18
Gcr1p	18
Gcr2p	20
Rap1p	21
Abf1p	23
Reb1p	25
Gcr1p as a DNA-Binding Protein	26
MATERIALS AND METHODS	27
Bacterial Strains	27
Media and Growth Conditions	27
Transformations	27
Induction of *Mal::GCR1* Gene Fusions	27
Purification of MBP-Gcr1p Fusion Proteins	28
Nucleic Acid Manipulations	30
DNA Precipitation	30
Purification of DNA Fragments and Oligonucleotides	31
Radio-Labeling of DNA Fragments	31
Polyacrylamide Gel Electrophoresis	34
Determination of DNA Concentrations	35
Generation of Double Strand DNA Oligonucleotides	35
DNA Sequencing	36
Plasmid Construction	37
In Vitro Transcription	42
In Vitro Translation	43
In Vitro DNase I Protection Assays	44
DNA Band-Shift Assays	45

Titration of DNA-Binding Activity 46
Calculation of Equilibrium Binding Constant 46

RESULTS 47

Gcr1p Expressed in *E. coli* or Rabbit Reticulocyte Lysate Binds to DNA 48
The DNA-Binding Domain of Gcr1p Resides within the Carboxy-Terminal 154 Amino Acid Residues 52
DNaseI Footprint Analysis 57
Identification of a Consensus Gcr1p Binding Sequence 65
The Gcr1p-DNA Nucleoprotein Complex 68
The DNA-binding affinity of the GCR1 binding domain 68
Specificity of the Gcr1p DNA-Binding Domain 77
Bending of DNA in the Gcr1p-DNA Nucleoprotein complex 81

DISCUSSION 89

The DNA-Binding Domain of Gcr1p 92
A DNA Consensus Sequence for Gcr1p 95
The Binding Affinity of the Gcr1p DNA-Binding Domain 97
The Bending of DNA by Rap1p and Gcr1p 106
The Significance of Adjacent Rap1p and Gcr1p Binding Sites in the UAS of the Glycolytic Genes 108
Conclusions 113

REFERENCES 115

BIOGRAPHICAL SKETCH 139

LIST OF TABLES

Table		page
1.	Oligonucleotides	32
2.	Plasmids	38
3.	Oligonucleotides containing CTTCC sequences	40
4.	DNA-Binding Affinity of Select Transcriptional Activators	99

LIST OF FIGURES

Figure		page
1.	Comparison of DNA-binding activity of Gcr1 and hybrid MBP-Gcr1p fusion protein	51
2.	Templets used to generate carboxy-terminal deletions of Gcr1p	53
3.	Carboxy-terminal truncation polypeptides of Gcr1p	55
4.	DNA-binding activity of Gcr1p carboxy-terminal deletion polypeptides	56
5.	*malE::GCR1* gene fusions	58
6.	DNA-binding activity of MBP-Gcr1p hybrid proteins	60
7.	Summary of data mapping the Gcr1p DNA-binding domain	61
8.	Gcr1p DNA-binding domain protects the CTTCC sequence motif in UAS_{TPI1}	64
9.	Gcr1p DNA-binding domain binds to CTTCC sequence elements found in front of other glycolytic genes	67
10.	Titration of Gcr1p DNA-binding activity	72
11.	Determination of Gcr1p DNA-binding domain binding affinity	75
12.	Graphical representation of binding affinity ...	76
13.	Competition experiment using specific competitor	78
14.	Competition experiment using non-specific competitor	79
15.	Graphical representation of competition experiments	80
16.	Circular permutation assay	83

17. Schematic of probe used in circular permutation assay 85

18. DNA is bent in the Rap1p-DNA nucleoprotein complex 86

19. DNA is bent in the Gcr1p-DNA nucleoprotein complex 87

20. SAPS analysis of Gcr1p 94

Abstract of Dissertation Presented to the Graduate School of the University of Florida in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

CHARACTERIZATION OF THE DNA-BINDING ACTIVITY OF THE *SACCHAROMYCES CEREVISIAE* TRANSCRIPTIONAL ACTIVATOR GCR1P

By

Michael Andrew Huie

August 1994

Chairman: Henry V. Baker, Ph.D.
Major Department: Immunology and Medical Microbiology

The enzymes of the glycolytic pathway constitute approximately 50 percent of the soluble proteins of the yeast *Saccharomyces cerevisiae*. Deletion of the gene encoding the transcriptional activator Gcr1p results in a 20-fold reduction of these enzymes. This study presents a biochemical analysis of the DNA-binding activity of Gcr1p. The DNA-binding domain of Gcr1p is mapped to the carboxy-terminal 154 amino acids of the polypeptide. DNase I protection studies presented here show that the Gcr1p DNA-binding domain protects a region of the upstream activating sequence of *TPI1* harboring the CTTCC sequence motif. This sequence has been shown by genetic methods to be important for high-level gene expression of a number of the glycolytic enzymes. By DNA band-shift assays it is shown that the Gcr1p DNA-binding domain also forms nucleoprotein complexes with CTTCC sequence elements found in the upstream activating sequences of *PGK1*,

ENO1, *PYK1*, and *ADH1*. From these experiments a consensus Gcr1p-binding site is derived which is 5'-(T/A)N(T/C)N(G/A)NC(T/A)TCC(T/A)N(T/A)(T/A)(T/G)-3'. The apparent dissociation constant of the Gcr1p DNA-binding domain with the sequence 5'-TTTCAGCTTCCTCTAT-3' is 2.9×10^{-10}M. However, only a 33-fold difference is observed between the ability of specific competitor and random DNA to inhibit formation of the nucleoprotein complexes between Gcr1p and this binding site. Circular permutation DNA band-shift assays are used to show that the Gcr1p-DNA nucleoprotein complexes contains bent DNA. The implications of these findings, in terms of the combinatorial interactions that occur at upstream activating sequences of *GCR1*-dependent genes, are discussed.

INTRODUCTION

> We can, first, describe an organism with concepts men have developed through contact with living beings over the millennia. In that case, we speak of living, organic function, metabolism, breathing, healing, etc. Or else we can inquire into causal processes. Then we use the language of physics and chemistry, study chemical or electrical processes, and assume, apparently with great success, that the laws of physics and chemistry, or more generally the laws of quantum theory, are fully applicable to living organisms. These two ways of looking at things are contradictory. For in the first case we assume that an event is determined by the purpose it serves, by its goal. In the second case we believe that an event is determined by its immediate predecessor. It seems most unlikely that both approaches should have led to the same result by pure chance. In fact, they complement each other, and, as we have long since realized, both are correct precisely because there is such a thing as life. Biology thus has no need to ask which of the two viewpoints is the more correct, but only how nature managed to arrange things so that the two should fit together.
>
> Niels Bohr (in Heisenberg, p. 110)

The regulation of gene expression in space and time is foundational to the generation of structural and functional diversity in living systems. It is in regulation that the abstract information content of DNA is made manifest and interactive with the environment. Using the working metaphor of the genetic code as the language of life, to a large degree the alphabet has been determined. The next challenge in molecular biology is the determination of the operational

syntax. Because as in language this allows unlimited arrangement of a finite set to generate creativity and innovation, an understanding of these rules will lead to an insight of life currently unavailable.

It has been pointed out long ago (Mayr, 1970; Britten and Davidson, 1969) that most structural genes do not differ greatly between simple single-cellular organisms and their more sophisticated mammalian brethren. The tremendous difference in complexity between the two forms of life seems to be due chiefly to the emergence of algorithms able to generate new biological structures with novel function. In fact, analysis of the rates of mutation, using either neutral or selection theory, suggest that mutations are insufficient to drive saltatory speciation unless they affect regulatory genes (King and Wilson, 1975; Delbrück, 1975; Gould, 1977; Jacob, 1977). Stated by the neo-Darwinian evolutionist Ernst Mayr:

> The fact that the macromolecules of most important structural genes have remained so similar, from bacteria to the highest organism, can be much better understood if we ascribe to the regulatory genes a major role in evolution. Since they strongly affect the viability of the individual they will be major targets of natural selection... The day will come when much of population genetics will have to be rewritten in terms of the interaction between regulator and structural genes. (Mayr, 1970, p.183)

Indeed, developmental biology, molecular genetics, cellular biology, evolutionary biology, and physical biochemistry have been converging toward a unified theory in recent years. As stated recently in an editorial justifying the existence of a new journal: Virtually all of the biological sciences are

converging on two central phenomena, namely, signal transmission and differential gene expression (Levine, 1989).

The physical structure of DNA and proteins places constraints on the available mechanisms of transcriptional regulation, and this is seen in the strong evolutionary conservation of many aspects of the transcriptional machinery (Guarente and Bermingham-McDonogh, 1992). This provides a tremendous advantage in the research of eukaryotic gene regulation because many processes can be dissected in the yeast *Saccharomyces cerevisiae*, using the powerful methods available for the study of this organism. Not surprisingly many recent advances have come from the study of *S. cerevisiae*.

In the complicated area of the regulation and initiation of transcription by RNA polymerase II, a large number of observations and experiments have lead to a heuristic model. Many components involved in promoter specific activation, the basal transcriptional machinery, chromatin structure, and nuclear scaffold have recently been cloned and sequenced. Biochemical and genetic analysis has revealed relationships and interactions among these factors (reviewed in Gasser and Laemmli,1987; Workman and Buchman, 1993). Factors involved in communication between activation domains and the basal transcriptional machinery have been isolated (Dynlacht et al.,1991; Berger et al.,1990 and 1992; Flanagan et al., 1991). I will present a current view of transcriptional activation based on recent information and then relate it to

a specific system, that of glycolytic gene expression in *S. cerevisiae*.

Basal Transcriptional Machinery and TFIID

Transcriptional Factor IID (TFIID) has long been believed to be the key link between promoter-specific activation and RNA polymerase II basal initiation machinery. TFIID is not a single protein but a large (>700 Kd) complex (Pugh and Tjian, 1990; Dynlacht et al., 1991) which elutes at 0.6-1.0 M KCl from a phosphocellulose column of human tissue culture cells (Matsui et al.,1980; Samuels et al., 1982; Davison et al., 1983). Reconstitution experiments had revealed that TFIID is the only component capable of sequence-specific DNA binding (Sawadogo and Roeder, 1985; VanDyke et al.,1989; Davison et al., 1983). Largely due to this observation and similar reconstitution experiments (Buratowski et al., 1988), TFIID was proposed to be the first factor to bind DNA with the aid of TFIIA (Reinberg et al., 1987) and TFIIJ (Cortes et al., 1992) and then to recruit the other basal machinery (TFII,-B,-E,-F,-G, and RNA polymerase II) (reviewed in Saltzman and Weinmann, 1989; Roeder, 1991). Additionally, *in vitro* the stable association of TFIID with the TATA box is slow (Reinberg and Roeder, 1987; Schmidt et al., 1989; Hahn et al., 1989; Lieberman et al., 1991). From these observations the association of TFIID with the TATA box is believed to be the first and rate limiting step in transcription initiation. Presumably, promoter-specific activators speed up or stabilize TFIID interaction with the TATA box thereby allowing recruitment of the other basal

machinery, thus stimulating transcription (reviewed in Ptashne, 1988 and Ptashne and Gann, 1990). This is not the only view put forward, however. Lin and Green (1991) have proposed that binding of TFIIB represents the rate limiting step that is enhanced by activators. The possibility that different activators target different steps in assembly is certainly reasonable (Mitchell and Tjian, 1989; Hawley, 1991).

Characterization of TFIID had remained limited due to the fact that it is difficult to purify in large yield and due to its instability. Recently, however, it was found that yeast TFIID was functionally interchangeable with human TFIID (Buratowski et al., 1988; Cavallini et al., 1988). Easier purification of the yeast factor was followed by molecular cloning of the TATA-binding protein (TBP) component of the complex, in yeast (Hahn et al., 1989; Eisenmann et al., 1989; Horikoshi et al., 1989; Schmidt et al., 1989; Cavallini et al., 1989). In yeast the 29 Kd TBP appears to be the major and perhaps single component of TFIID. Cloning of TBP from human (Peterson et al., 1990; Hoffmann et al., 1990) and other species (Fikes et al., 1990; Hoey et al., 1990; Gasch et al., 1990; Haass et al., 1992; Kao et al., 1990; Tamura et al., 1991; Muhich et al., 1990; Ganster et al., 1991; Wong et al., 1992) soon followed due to the remarkably high degree of homology of 180 amino acid primary sequence at the carboxyl-terminus of the protein.

TBP is composed of a conserved carboxy-terminal domain and a divergent amino-terminal domain. Sequence analysis

revealed that TBP is a relatively small protein varying amongst species between 22-39 Kd. The amino acids in the carboxy-terminus of TBP are 75-85% identical in all eukaryote TBP sequences currently known (Ganster et al., 1993; McAndrew et al, 1993). This region, called the C-terminal core domain, contains two repeats of 66-67 amino acids separated by a highly basic region. The structure of TBP from *Arabidopsis thaliana* has been determined by X-ray diffraction crystallography to a resolution of 2.6 Å (Nikolov et al., 1992). The protein folds into two symmetrical and topologically identical domains each derived from one of the repeats (Nikolov et al., 1992; Rigby, 1993). The presumed DNA binding surface is a curved, anti-parallel ß-sheet resembling a "saddle." Computer modeling revealed that DNA would fit nicely into the concave surface of the saddle (Nikolov et al., 1992; Rigby, 1993). When bound to DNA the convex surface of the saddle is able to interact with other transcriptional factors. Mutations in TBP that affect DNA binding map to the concave region, while mutations that affect the ability of TBP to interact with other proteins map to the convex surface (Rigby, 1993).

The cloning of TBP along with the cloning of the major components of TFIIB (Malik et al., 1991), TFIIE (Peterson et al., 1991; Ohkuma et al., 1991; Sumimoto et al., 1991), and TFIIF/RAP70/RAP30 (Aso et al., 1992; Finkelstein et al., 1992) has allowed studies of the interaction between these factors and activators to ensue. Recently, it has been shown that TBP is also required by RNA polymerases -I and -III for

specific initiation (Comai et al., 1991; Schultz et al., 1992; White et al., 1992).

Activators, Coactivators, and Adaptors

Sequence-specific transcriptional factors are comprised of two critical and separable domains: a DNA-binding domain and an activation domain. This organization was first demonstrated by Brent and Ptashne (1985) in the "domain-swapping" experiments with LexA and Gal4p. When the DNA-binding domain of LexA was fused with the activation domain of Gal4p, transcriptional activation occurred through a LexA-binding site. Even more dramatic, the DNA-binding, ligand-binding, and activation domains of the estrogen and glucocorticoid receptors could be swapped to produce an estrogen responsive glucocorticoid receptor (Green and Chambon, 1987); or the ligand binding and activation domain of the glucocorticoid receptor could be fused to Gal4p to produce hormonal dependent activation at a Gal4p-binding site (Hollenberg and Evans, 1988). This flexibility with which domains can be swapped has been one of the many surprises in the study of transcriptional activating proteins.

Another surprise was that often the secondary structure of these domains are not pre-folded but are formed and stabilized upon ligand interaction. Although the DNA-binding domains often have well defined motifs, ligand binding often contributes to their stability. For example, the DNA-binding domain of Gcn4p undergoes a coil to helix transition upon binding to its operator site (Weiss et al., 1990; Talanian et al., 1990; O'Neil et al., 1990). This loosely structured

domain apparently wraps around the DNA helix upon binding. A similar occurrence is also seen in the N-terminal arm of the lambda repressor which is disordered in the free protein but wraps around the 'back side' of the DNA helix when bound (Jordan and Pabo, 1988).

Activation domains appear to become structured upon interaction with other molecules. By selecting for activation from random *E. coli* genomic sequences Ma and Ptashne (1987) showed that a high density of negative charges, rather than well-defined amino acid sequences, was the major determinant of an activation domain in yeast. This type of activation domain has been termed an "acid blob" or "negative noodle" and has been proposed to be disorganized (Sigler 1988), an alpha helix (Irwin and Ptashne, 1987), or a beta sheet (Leuther et al., 1993; Van Hoy et al., 1993). Polypeptides of the acidic activation domain of Gal4p and Gcn4p have been shown to form beta sheets by circular dichroism spectroscopy (Van Hoy et al., 1993) and by genetic studies (Leuther et al., 1993). However, the active form of the activation domain may change with additional protein interactions (Hahn, 1993).

Another type of activation domain has been described and also displays a "disorganized" structure. Deletion analysis has revealed that the most potent activation domain of Sp1 contains ~25 percent glutamine and very few charged amino acids (Courey and Tjian, 1988). Furthermore, a glutamine-rich stretch of 145 amino acids from *Antennapedia* can substitute for this activation domain in a swap experiment

(Mitchell and Tjian, 1989). Besides the glutamine content there is no other obvious sequence homologies between these interchangeable domains. Finally, a third type of activation domain, mapped in CTF/NF1, is proline-rich (Mermod et al., 1989).

Various models have been proposed for how activation domains work. The models emphasize contact between the activation domain and TFIID or RNA polymerase II. For example, the negative noodle activation domain has been proposed (Sigler, 1988; Allison et al., 1988) to interact with the carboxy terminal domain (CTD) repeat of the large subunit of RNA polymerase II. A general feature of the large subunit of eukaryotic RNA polymerase II is the multiple repeat of the Tyr-Ser-Pro-Thr-Ser-Pro-Ser heptapeptide in the carboxyl terminus. In yeast this is repeated 26 times (Allison et al., 1985); in mouse, 52 times (Corden et al., 1985). Except for proline, which is believed to stabilize a unique secondary structure, nearly all the amino acid side chains in the heptapeptide are hydrophilic. This domain is believed to be fully exposed to solvent projecting out of the remaining globular folded polypeptide. According to the acid blob/negative noodle model, interactions between the carboxylates of the noodle and the hydroxyl groups of the CTD and other promoter binding proteins facilitate transcriptional initiation, possibly by aiding in binding of RNA polymerase II to DNA (Suzuki, 1990). TFIID interacts with the CTD repeats implying that its presence is necessary in this model (Sharp, 1992).

An alternative explanation of the role of acidic residues in activation domains has been proposed (Leuther et al., 1993) in which the acidic residues function to make the hydrophobic residues more accessible to interact with other factors. According to this hypothesis positive as well as negative charged amino acids can function in activation domains, and the hydrophobic residues are more important in making contacts with other proteins. Detailed mutagenesis studies of some acidic activation domains support this proposal (Cress and Triezenberg, 1991; Hardwick et al., 1992; Regier et al., 1993; Leuther et al., 1993).

The idea that these different types of activation domains all interact directly with TBP seemed unlikely, and indeed, after the cloning of TBP, various investigators reported that this protein alone could not replace the TFIID fraction in transcriptional activation by Sp1 or CTF (Pough and Tjian, 1990; Peterson et al., 1990), NTF1 (Dynlacht et al., 1991), or USF (Hoffmann et al., 1990). Furthermore, these activators apparently do not stimulate binding of TBP to the TATA- box by direct interaction. Proteins that carry out this function have been termed coactivators or adaptors/ mediators (reviewed in Lewin, 1990; Martin 1991;and Greenblatt, 1991). A large number of polypeptides can be coimmunoprecipitated with antibodies to TBP and are referred to as TAFs (for TBP associated factors) (Dynlacht et al., 1991). Recently, by biochemical and genetic methods, a number of RNA polymerase II TAFs (Taggart et al., 1992; Ganster et al., 1993; Eisenmann et al., 1992; Koleske et al.,

1992) and RNA polymerase III TAFs (Buratowski and Zhou, 1992; Kassavetis et al., 1992) have been isolated. These, apparently species specific, proteins are believed to be part of the heterogenous TFIID complex.

By deletion analysis some of these adaptor domains have already been mapped. For example, the acidic activator Gal4p-VP16 is able to activate transcription in a heat-treated TFIID deficient HeLa nuclear extract when supplemented with recombinant human TBP. However, a amino-terminally truncated molecule of human TBP cannot support the activation in heat treated HeLa nuclear extract (Peterson et al., 1990). Also, by genetic methods several putative adaptors have been isolated in yeast (Berger et al., 1992).

History of Fermentation in Yeast

Recent studies on the regulation of the glycolytic enzyme genes in yeast suggest candidates for activators and adaptors. I will first review the history of the study of fermentation in yeast since it is intimately associated with the history of modern biology. Then I will outline current factors imputed in transcriptional regulation of the glycolytic genes. In the discussion I will present a model based on current knowledge.

The brewing of beer is the largest biotechnological industry in the world producing commercially each year 10^{11} liters of beer (Oliver, 1991). The earliest known remains of human writing include descriptions of alcohol production by fermentation. Clay tablet found from the earliest

civilization, dating to 2,500 BC, written in the Sumerian and Akkadian languages describe an established profession of alcohol fermentation 3,500 years earlier. Assyrian writings dating to 2,000 BC list beer as a commodity on Noah's Ark. The Bible explicitly mentions Noah's knowledge of fermentation and his use of alcohol. (Genesis 9:20-21: Noah was the first tiller of the soil. He planted a vineyard; and he drank of the wine, and became drunk, and lay uncovered in his tent.) The daughters of Lott (Genesis 20:30-37) found creative use for alcohol. Egyptian writings from the Fourth Dynasty, *circa* 2,500 BC, describe the process of malting barley and its fermentation. Man's knowledge of these techniques clearly preceded his acquisition of written language. The effects of the products of fermentation, principally ethanol, undoubtedly influenced the perceptions and analysis of many of the unknown artists and religious charismatics of the antediluvian past. Anthropological studies reveal that all cultures independently learned early to ferment fruits and grains to produce alcohol.

The scientific study of fermentation by yeast has spanned many centuries and has been associated with and attributed with changes in fundamental beliefs in modern biology. Indeed, historians tell us that the study of fermentation in yeast underlies the *birth* of chemistry, biochemistry, and microbiology (Fruton, 1976; Kohler, 1972). Speculation on the nature of fermentation was advanced by some of history's most notable scientists.

Willis (1659) and Stahl (1697), founders of the

phlogiston theory of combustion, were the first to outline a theory of fermentation in the early modern period (for an excellent review, see Fruton, 1972). They attributed an internal motion to the "ferment." Contact of quiescent substances by the ferment then caused them to undergo decay. A similar view was also put forth by Isaac Newton (1730), in his book *Opticks*, where he describes the interaction of ferments as an example of the force of gravity.

Lavoisier, the founder of modern chemistry, extinguished the phlogiston theory with his work describing oxygen. But it is in Lavoisier writings on fermentation, not oxygen, where he outlines his theory of conservation of matter:

> Nothing is created either in the operation of art or of nature, and the principle may be formulated that in every process there is an equal quantity of matter before and after the reaction, that the quality and the quantity of the principles are the same and that there are only changes or modifications. (in Fruton,1972, p. 39)

Lavoisier, and later Gay-Lassic (1810), used fermentation to demonstrate this 'conservation of matter.' This 'law,' in its form modified by Einstein, remains a foundation of modern chemistry and physics. Although the phlogiston theory of combustion was abandoned, Lavoisier was largely responsible for sustaining the 'purely chemical' theory of fermentation that dominated the scientific opinion during his day.

In 1779, the French Academy of Science offered the prize of one kilogram of gold to anyone who could explain the nature of alcoholic fermentation (Fruton, 1972). The question was stated as, 'What are the characteristics which distinguish vegetable and animal substance that act as

ferments from those that undergo fermentation?' Offering this prize reflected not only the importance of this unsolved problem, and its relation to commerce, but the fact that beginning around the nineteenth century chemical explanations for biological phenomena became a major preoccupation of leading scientists (Fruton, 1972). A debate ensued as to whether vitalistic or organismic notions, beyond materialism and reductionism, had to be evoked to account for fermentation. The prize was never awarded, and was withdrawn in 1804 because of lack of funds.

In 1838, the same period that Schleiden and Schwann (1839) were outlining the cell theory, Cagniard de Latour published an article entitled Memoir on Vinous Fermentation (translated in Williams and Steffens, 1978) in which he argued that fermentation was the result of the 'vital activity' of yeast 'cells.' Although Leeuwenhoek, in 1680, described multiplying budding cells in the deposit formed in beer fermentation, it is important to remember that during this time, fermentation was still believed to be a chemical (that is, *not* vitalistic or organismic) process, and that yeast sediment was not believed to be cellular life, but an albuminoid. Cagniard de Latour discovered that yeast sediment was composed of cells and that fermentation was associated with these cells being alive and dividing (Cagniard-Latour 1838). He also immediately recognized the importance of these new cellular organisms as a research tool, as shown in the concluding remarks of his paper:

> I have looked at the principal works which treat vinous fermentation and in none of them have I seen

> that the microscope has been used to study the phenomena upon which it depends. This essay, as one can judge by the researches which I have just mentioned, will be useful since it has furnished a number of new observations from which the principal results that can be drawn are: 1, that beer yeast, this ferment which is in such use and which, for this reason, should be examined very closely, is a mass of little globular bodies capable of reproducing themselves, and thus organized beings, and not a simple organic or chemical substance as has been supposed; 2, that these bodies seem to belong to the vegetable kingdom and regenerate themselves in two different ways, and 3, that they seem to act on a solution of sugar only so long as they are in the vital state: from which it can be concluded that it is very probably by some effect of their vegetation that they are able to disengage carbon dioxide from this solution and convert the solution into a spirituous liquor. I would like to remark, further, that yeast, considered as an organized matter, perhaps merits the attention of physiologists in this sense: 1, that it can be born and develop in certain circumstances with great rapidity even in the middle of carbonic acid as in the brewers' barrels; 2, that its mode of reproduction presents particularities of a kind which have not been observed in other microscopical products composed of isolated globules and 3, that it does not die by very considerable refrigeration nor by deprivation of water. (in Williams and Steffens, 1978,p. 446)

Cagniard de Latour's work was confirmed by Theodore Schwann (1838), yet it still was attacked and dismissed by many of the scientists of the day who maintained institutional power, such as Berzelius, Liebig, and Wohler. This group of prominent scientists were chemist who had succeeded at synthesizing bio-organic compounds and found the notion of 'vital forces' as somewhat superstitious. 'It is not,' said Liebig, 'because it is organized that the beer yeast is active, but because it has been in contact with the air. It is the dead portion of the yeast, which has been alive and is in process of alteration, that acts on the sugar.' (in Pasteur Valler-Radot, 1957)

Pasteur, who was held in respect among chemists, supported the work of Cagniard de Latour and Schwann, and was instrumental in countering the opposition of the chemists, and in bolstering the concept of vitalism in biology. Said Pasteur:

> My present and most fixed opinion regarding the nature of alcoholic fermentation is this: I believe that there is never any alcoholic fermentation without there being simultaneously the organization, development, and multiplication of the globules,or at least the pursued, continued life of globules that are already present. The totality of the results in this article seem to me to be in complete opposition to the opinions of Liebig and Wohler. (in Kornberg,1989, p. 33)

Pasteur strongly supported an ideology vitalism and held that life processes are not reducible to the laws of physics and chemistry.

Pasteur's 'fixed opinion' was shown to be false when Eduard Buchner made the following observations in 1897:

> If one mixes 1,000 grams of brewer's yeast with an equal weight of sand and then grinds the mixture, the mass becomes moist and pliable. Now if 100 grams of water are added, and the paste, wrapped in cheesecloth, is gradually subjected to 400-500 atmospheres of pressure in a hydraulic press, one obtains about 500 milliliters of 'press juice'; to remove any residual unbroken cells, the press juice is passed through a paper filter. The final solution contains a collection of substances derived from the cell interior. The 'cell extract' obtained in this way is a clear, slightly yellow liquid with a pleasant yeast smell. (in Kornberg, 1989, p. 34)

When Buchner added sugar to this solution in an effort to preserve it, he noticed that bubbles of gas appeared within minutes. He repeated his experiments and concluded that fermentation was possible outside of the yeast cell. Buchner believed that the reaction was catalyzed by a single enzyme and named the substance zymase. The word enzyme means 'in

yeast.' Buchner received the Nobel Prize in Chemistry ten years later for his discovery, which is often cited as the *origin* of biochemistry. In his autobiography Willstatter, Buchner's teacher, said 'This will bring him fame, even though he has no chemical talent.' (in Willstatter, 1949)

Studies of Pasteur's records revealed that he had prepared cell-free yeast juices and attempted to carry out fermentation with the juice before Buchner's discovery. Unfortunately, he had used a strain of yeast which contained a labile form of invertase which did not survive the extraction procedure. He did not observe cell-free fermentation (Kornberg, 1989).

Early work on the purification of enzymes then ensued. Although this early work was criticized on the ground that it was 'unphysiological to separate enzymes from cell,' investigators persevered and the conception of enzymes changed from a vague property in certain preparations to definite chemical substances, and finally by the 1930s to specific proteins (Fruton, 1972). The work of Harden, Neuberg, Embden, Meyerhof, and Warburg showed that zymase was a complex mixture of a dozen separate enzymes.

In the latter half of the twentieth century, after the revolution in molecular biology, the genes coding for the enzymes of the glycolytic pathway were cloned and the study of transcriptional regulation of these gene ensued.

Coordinate Regulation of Glycolysis in *Saccharomyces cerevisiae*

In *Saccharomyces cerevisiae* the enzymes of the glycolytic pathway are among the most highly expressed genes, constituting between 30-60 percent of the total soluble protein when grown in the presence of glucose (Hess et al., 1969; Fraenkel, 1982). The demonstration (Holland and Holland, 1978) that the most abundant mRNA species in the yeast code for glycolytic enzymes suggested that this is largely due to high-level transcription of the corresponding genes. High-level expression of heterologous genes under glycolytic-gene promoters (Bitter and Egan, 1984; 1988; Bitter et al., 1987) further demonstrated that the glycolytic gene promoters are among the strongest known.

A number of *trans*- and *cis*-acting elements have been discovered to be important for the coordinate high-level expression of the glycolytic genes. The *trans*-factors include Gcr1p, Gcr2p, and Rap1p/Grf1p/Tuf1p (hereafter referred to as Rap1p). Additionally the factors GAL11/SPT13 (referred to as Gal11p), Abf1p/Taf1p/Sufp/Gf1p/SBF-B (now designated Abf1p) and Reb1p/Grf2p/QBP/Y (hereafter referred to as Reb1p) are believed to play a role in high-level expression of some of the genes. I will review these factors here. A model of how these factors may work together will be presented in the Discussion.

Gcr1p

It was observed in 1978 that mutation in a single gene can causes a dramatic decrease in expression of most of the glycolytic genes (Clifton et al., 1978; Clifton and Fraenkel,

1981). This gene was named *GCR1* for Glycolysis Regulation. The *gcr1* mutant has a severe growth defect when glucose is present in the medium. This growth defect is presumably due to the fact that aerobic respiration is repressed in *Saccharomyces cerevisiae* when glucose is present, as part of the global phenomenon known as glucose repression. When glucose is present a large number of genes involved in the metabolism of alternate carbon sources are repressed. Genetic analysis has identified a large number of genes involved in this regulatory process, involving multiple steps and branches regulating subsets of glucose-repressible genes (for a recent review see Trumbly, 1992). Since the *gcr1* mutant produces the glycolytic enzymes in reduced amounts it grows poorly in the presence of glucose. It does, however, grow adequately on non-carbohydrate carbon sources, and if glucose is added a noticeable induction of most of the glycolytic enzymes is observed (Clifton and Fraenkel, 1981; Baker, 1986). *GCR1* has been cloned by complementation (Kawasaki and Fraenkel, 1982), mapped to chromosome XVI (referred to as the *sit3* mutant in Arndt et al., 1989, and Devlin et al., 1991), and sequenced (Baker, 1986; Holland et al., 1987) revealing an open reading frame coding for a protein of 844 amino acids. The low codon bias of -0.00086 according to the rules of Bennetzen and Hall (1982) suggested that the gene is expressed in low amounts, and this was confirmed by Northern analysis (Baker, 1986). A single base pair insertion at codon 304 causing a frame-shift mutation is apparently responsible for the phenotype associated with the

gcr1-1 mutant strain DFY67 (Holland et al., 1987). Sequence analysis suggested that a possible helix-turn-helix (H-T-H) motif is present in the carboxyl terminal region of the protein (Baker, 1986). H-T-H motifs are associated with DNA-binding activity. Gcr1p was shown to posses DNA-binding activity when it was demonstrated that the Gcr1p product translated in rabbit reticulocyte lysate formed a nucleo-protein complex with a DNA fragment isolated from the upstream activation sequence element of *TPI1* (Baker, 1991).

Gcr1p was shown to bind a fragment of DNA from the *TPI1* promoter which contained a CTTCC pentamer but was unable to bind a related fragment in which the CTTCC sequence was changed to CAACC (Baker, 1991). The CTTCC sequence element had long been noted to be present in the promoters of glycolytic genes (Ogden et al., 1986). By site-directed mutagenesis or deletion analysis CTTCC sequence elements were shown to be important for high-level expression of the genes encoding phosphoglycerate kinase (Chambers et al., 1988), enolase and pyruvate kinase (Buchman et al., 1988), and triose phosphate isomerase (Scott and Baker, 1993). Bitter et al. (1991) defined a sequence, GPE, which had a CTTCC sequence element at its core in upstream activation sequences.

Gcr2p

Mutation or deletion in the *GCR2* locus also produces a profound effect on expression of the glycolytic genes. *gcr2* mutants were isolated during a screen for mutants affecting expression of a *ENO1::lacZ* gene fusion (Uemura and Fraenkel,

1990). The *GCR2* gene was cloned by complementation and shown to be distinct from *GCR1*. Although the effects on glycolytic gene expression are similar in the *gcr1* and *gcr2* mutants, the *gcr2* mutant does not have as severe of a growth defect as *gcr1* mutants when grown on glucose at 30 C (Uemura and Fraenkel, 1990).

The role of Gcr2p in glycolytic gene expression is currently under investigation. Like Gcr1p, Gcr2p appears to exert its effects at the transcriptional level, as measured by Northern analysis (Uemura and Fraenkel, 1990). Gcr2p does not appear to interact directly with the CTTCC sequence element. *In vivo* footprinting experiments demonstrate protection of the CTTCC sequence elements in UAS_{TPI} in the *gcr2* mutant (Scott and Baker, 1992). In the *gcr1* mutant these sequences are not protected (Huie et al., 1991; Scott and Baker,1993). A clue to the function of Gcr2p in glycolytic gene expression was provided by genetic studies (Uemura and Jugami,1992) based on the method of Fields and Song (1988). A *GAL4::GCR1* gene fusion containing the activation domain of GAL4 and amino acids 68-844 of Gcr1p can complement a *gcr2* mutant. Furthermore, a *RAP1::GCR2* gene fusion which carries the DNA binding domain of Rap1p can partially complement a *gcr1* mutant. These observations have led to the suggestion that Gcr1p and Gcr2p function together as an activation complex (Uemura and Jigami, 1992).

Rap1p

RAP1, for repressor/activator protein, is an essential gene (Shore and Nasmyth, 1987) whose product binds to DNA.

The *RAP1* product is expressed at levels greater than 4,000 molecules per cell (Buchman et al., 1988) and has pleiotropic actions. It is capable of either activation or repression of transcription depending on the context of its binding site (Shore and Nasmyth, 1987). For example, Rap1p acts as a silencer at the mating-type locus by interacting with the HMR(E) element (Shore et al., 1987). Rap1p acts as an activator of ribosomal protein genes in cells in exponential growth by interacting with the *rpg*-box found upstream of the ribosomal protein genes (reviewed in Warner, 1989). It is also involved in stringent control of ribosomal protein transcription under conditions of amino acid starvation (Moehle and Hinnebusch, 1991). Rap1p binds to telomeres (Buchman et al., 1988; Longtine et al., 1989) and is involved with their maintenance and length regulation (Lustig et al., 1990; Conrad et al., 1990). Interestingly, Sussel and Shore (1991) showed that two of these three functions can be separated by genetic methods. They were able to distinguish Silencing/Telomere function from activation function. A DNA-binding domain (Henry et al., 1990) and activation and derepression domains (Hardy et al., 1992) have been mapped within the Rap1p protein.

A role for Rap1p in the activation of the glycolytic genes has been demonstrated for a number of genes including *TPI1* (Scott et al., 1990), *TDH3* (Bitter et al., 1991), *PGK* (Chambers et al., 1989), *ENO1* (Machida et al., 1989; Brindle et al., 1990), *ENO2* (Brindle et al., 1990), *PYK* (McNeil et al., 1990), *PDC1* (Butler et al., 1990), and *ADH1* (Tornow and

Santangelo, 1990). Furthermore, putative Rap1p-binding sites can be identified in the 5' region of other genes in the glycolytic pathway (Huie et al., 1992; and discussion). Rap1p appears to undergo phosphorylation which may influence its binding at the UAS_{PGK} (Tsang et al., 1990). This post-translational modification has been suggested (Tsang et al., 1990) to be responsible for the glucose induction of *PGK* (Chambers et al., 1989).

The mechanism of action of Rap1p is currently unknown. Although Rap1p binding sites are necessary for high-level expression of glycolytic genes, they are unable to convey high-level expression of reporter genes by themselves (Stanway et al., 1989). Protein-protein interactions with the putative coactivator GAL11/SPT13 has been proposed (Nishizawa et al., 1990; Fassler and Winston, 1989; Stanway et al., 1993), and the mapping of an activation domain in Rap1p to 66 amino acids with a net negative charge of -12 (Hardy et al., 1992) also supports a coactivation model (Baker, 1991). Alternatively, Rap1p has been proposed to alter chromatin structure providing access to DNA by other activator proteins (Devlin et al., 1991; Sentenac and Vignais, 1987). Finally, Scott and Baker (1993) have proposed that Rap1p facilitates binding of Gcr1p by either a protein-protein interactions or by altering the topology of the DNA to which Gcr1p binds.

Abf1p

Abf1p (for ARS binding factor), like Rap1p, is also involved in multiple cellular functions including

transcriptional activation and initiation of DNA replication. Binding sites for Rap1p and Abf1p are both found at HMR(E). Abf1p binds to the promoters of the ribosomal-protein genes encoding L2A and L2B (Seta et al., 1990). Abf1p binds to the B3 element (Marahrens and Stillman, 1992) of the autonomously replicating sequences (ARS) and is important for initiation of replication. Replacement of the Abf1p-binding site with a Rap1p-binding site can restore high-level ARS activity (Marahrens and Stillman, 1992). The binding sites for these factors can also be exchanged in a gene transcribed br RNA polymerase II: an Abf1p-binding site can replace a Rap1p-binding site in the *ILV1* gene (Remacle and Holmberg, 1992). A functional Abf1p-binding site is found in the UAS of *PGK* and *PYK1* (Chambers et al., 1990).

The similarity of function of Rap1p and Abf1p is matched by an homology of 40% conserved amino acid over 60% of the primary sequence between the two proteins (Diffley and Stillman, 1989). This has lead to a model of action of these two proteins (Diffley and Stillman, 1989). The repressed state of HMR involves the formation of chromatin structure similar to that of heterochromatin (Nasmyth, 1982). Heterochromatin is located in the periphery of the nucleus associated with the nuclear lamina. The *SIR4* gene is also required for silencing (Rine and Herskowitz,1987). The presence of a lamina domain in the Sir4p protein suggests that Rap1p and Abf1p may play a role in altering chromatin structure by interacting with the nuclear scaffold (Marshall et al, 1987; Diffley and Stillman, 1989). Indeed Rap1p has

been shown to purify with the nuclear scaffold (Cardenas et al, 1990), as do ARS sequences (Amati and Gasser, 1988).

Reb1p

REB1 is an essential gene in *Saccharomyces cerevisiae* encoding a highly hydrophilic protein of 809 amino acids (Ju et al., 1990). The protein was originally isolated as a factor binding to the rRNA enhancer (Morrow et al., 1989 and 1990), but now appears to play a more global role in transcription. It is likely to be identical (Ju et al., 1990) to the proteins Y (Fedor et al., 1988), Grf2p (Chasman et al, 1990), and QBP (Brandl and Struhl, 1990). The Reb1p consensus binding site is present upstream of a number of genes transcribed by RNA polymerase II, including the highly expressed glycolytic genes *PYK* (Chasman et al., 1990), *TPI1* (Scott and Baker, 1993), TDH3 (Bitter et al., 1991) and ENO1 (Machida et al., 189). The exact role of Reb1p in the expression of genes transcribed br RNA polymerase II is currently unknown; however, it has been shown to effect chromatin structure (Fedor et al., 1988), and although Reb1p-binding sites have little effect on activation by themselves, they do potentiate nearby activators (Chasman et al., 1990; Holmberg and Remacle, personal communication). This synergistic effect is strongly distance dependent. This observation has lead to the suggestion that Reb1p exerts its effect by clearing chromatin of nucleosomes, or other obstacles, allowing activators easier access to bind their sites (Kornberg and Lorch, 1991).

Gcr1p as a DNA-Binding Protein

The experiments presented in this study were designed to characterize several basic properties of the DNA-binding activity of Gcr1p. Since Gcr1p is a transcriptional activator it should adhere to the modular nature of this class of proteins. The experiments described in this study set out to determine whether Gcr1p does have a modular design and a DNA binding domain. Deletion analysis of Gcr1p was used to address this point. Experiments also were designed to determine if the Gcr1p DNA-binding domain binds to DNA in a sequence specific manner and to define a consensus sequence to which it binds. This was studied by DNase I footprinting and DNA band-shift assays. Whether the CTTCC core sequence is sufficient to convey DNA binding, or if surrounding base pairs are important in the ligand interaction was also addressed in this study by DNA band-shift assays. Determination of affinity and specificity of binding by Gcr1p to a consensus sequence is presented. Finally, since many DNA-binding proteins have been shown to distort DNA when bound, Gcr1p was also assayed for this activity. The implications of the findings in this study are discussed in the context of the combinatorial nature of interactions of factors important for glycolytic gene expression.

MATERIAL AND METHODS

Bacterial Strains

E. coli strains used in this study were MC1061 (*hsdR, mcrB, araD139, Δ[araABC-leu]7679, ΔlacX74, galU, galK, rpsL, thi*) (Casadaban and Cohen, 1980); DH5α (*ϕ80dlacZΔM15, Δ[lacZYA-argF]U169, deoR, recA1, hsdR17, supE44, thi-1, hyrA96, relA1*) (Hanahan, 1983); and, TB1 (*ara Δ[lac ZYApro AB] , rpsL [ϕ80dlacZΔM15], hsdR*) (Johnston et al., 1986).

Media and Growth Conditions

E. coli strains used in this study were grown at 37 *C* in L broth (5 g yeast extract, 10 g tryptone, and 8 g NaCl per liter) with vigorous shaking. Strains harboring plasmids were grown in L broth containing 100 μg/ml ampicillin.

Transformations

E. coli strains MC1061 and TB1 were transformed with plasmid DNA by the low pH method of Enea et al. (1975). *E. coli* strain DH5α was transformed by the manufacturers recommended method (GIBCO BRL).

Induction of *malE::GCR1* Gene Fusions

E. coli strain TB1, harboring plasmids encoding *malE::GCR1* gene fusions, was used for the production of hybrid MBP-Gcr1p polypeptide. 500 ml cultures in 2 liter shake flasks were grown to an optical density of 0.5 (A_{600}). Then isopropyl-ß-D-thiogalactopyranoside (IPTG) was added to

a final concentration of 2 mM to induce expression of the *malE::GCR1* fusions genes. The cultures were then grown an additional 2 hours at 37 *C*, following which time they were harvested by centrifugation (4000 x *g* for 10 min at 4 *C*). The supernatant was discarded and the wet weight of the pellet was determined. The cells were suspended in 3 ml TEN lysis buffer (50 mM Tris [pH 8.0], 1 mM EDTA, 50 mM NaCl) per gram (pellet [wet weight]). The bacteria were lysed by passage through a French pressure cell at 20,000 lb/in^2 by the method of Clifton et al. (1978). Cellular debris was removed by centrifugation at 17,000 x *g* for 20 min at 4 *C*. The supernatant, containing the soluble cellular protein, was recovered for further use. Protein extracts prepared in this manner typically had a protein concentrations that ranged from 0.068 to 0.2 mg/ml, as determined by the method of Bradford (1976).

Purification of MBP-GCR1 Fusion Protein

The Maltose-binding moiety of the hybrid MBP-Gcr1p(690-844) fusion protein was utilized to purify the fusion protein by affinity chromatography over an amylose column. The amylose column was prepared in the following manner. A 15 ml slurry of amylose resin was allowed to settle in a 2.5 x 10 cm column. The height of the packed resin was 1 cm giving a total bed volume of 7 cm^3. The column was then washed with approximately 3 column volumes (25 ml).

Crude cellular lysates prepared as described above were diluted 1:5 with column buffer (10 mM phosphate, 0.5 M NaCl, 1 mM azide, 1 mM DTT, 1 mM EGTA) containing 0.25% Tween 20 to

yield a final volume of 400 ml. The diluted lysate was loaded onto the amylose column at the rate of 1 ml/min. Following the addition of the crude lysate to the column, the column was washed with 3 column volumes of column buffer containing 0.25% Tween 20. The column was then washed with 6 column volumes column buffer without Tween 20. The hybrid MBP-GCR1(690-844) fusion protein was then eluted from the column by passing 10 nM maltose over the column in column buffer. The fusion protein was collected in 3 ml aliquots. Typically 15 fractions were collected.

Aliquots of 20 µl of each fraction were subjected to SDS-PAGE to identify fractions containing fusion protein. 20 µl aliquots of each fraction were electrophoresed through a 10% SDS-PAGE. The resulting gel was stained with coomassie blue to visualize the protein. Typically fusion protein appeared across fractions 2 through 7. Fractions containing the fusion protein were pooled and concentrated by ultrafiltration through a low-adsorption, hydrophilic, [YM] membrane with a 30 kiloDalton size exclusion using a Centriprep-30 concentrator (Amicon) according to the manufacture's specifications. Pooled fractions were added to the Centriprep-30 concentrator and centrifuged for 5 min at 2600 rpm (5000 x *g*) in a Jouan CRF412 swinging-bucket benchtop centrifuge. The filtrate was decanted and the sample was centrifuged a second time for 10 min at 2600 rpm. The final volume was approximately 1 ml. This sock solution was confirmed by running 2 µL into 10% SDS-PAGE and protein visualized by staining with coomassie blue. Samples were

saved by storage at -70 *C*. A typical protein concentrations of such a prep was 0.35 mg/ml.

Nucleic Acid Manipulations

Standard techniques used throughout the course of this study are described in common reference manuals (Ausubel et al., 1989; Sambrook et al., 1989). Deviations from standard techniques are noted and described.

DNA Precipitation

Prior to ethanol precipitation ammonium acetate was added to the DNA solutions to a final concentration or 2.5 M. This was accomplished by adding an equal volume of 5 M ammonium acetate. The resulting volume was noted and then 2.5 volumes of absolute ethanol was added, mixed, and placed at -70 *C* for 10 min. Following which time the samples were centrifuged in a microcentrifuge at 12,000 x *g* at room temperature for 20 to 30 min. After centrifugation the supernatant was discarded and the pellet washed with 1 ml 70% ethanol. The resulting mixture was then centrifuged for 5 min at 12,000 x *g* in a microcentifuge at room temperature. Again the supernatant was discarded and then the pellet was dried in vacuum with a Speed Vac Concentrator (Savant Instruments Inc.).

In situations were DNA was to be treated with polynucleotide kinase, ammonium acetate was avoided. In these cases DNA was precipitated by adjusting solution to 0.3 M sodium acetate (pH 5.5). Typically 1/10 volume of 3 M sodium acetate was added followed by addition of 2.5 volumes

of ethanol. Samples were cooled, centrifuged, and dried as described above.

Purification of DNA Fragments and Oligonucleotides

Oligo-nucleotides used in this study were initially synthesized on an Applied Biosystems 380B DNA synthesizer by the University of Florida Interdisciplinary Center for Biotechnology Research. The sequence of the oligonucleotides used are listed in Table 1. In some cases oligonucleotides were purified by 10% PAGE. After DNA was visualized by ethidium bromide staining, gel slices were incubated at 37 *C* overnight in 3 volumes of 0.5 M ammonium acetate, 1 mM EDTA, and then precipitated the next morning by the addition of 2.5 volumes ethanol and centrifugation.

DNA was purified from acrylamide gels by transfer to DEAE paper. The excised gel was embedded in 0.8% agarose gel. Current was applied at constant voltage of 100 V and the DNA was run into NA45 DEAE cellulose paper (pre soaked for 10 min in 10 mM EDTA pH 7.6; then 5 min in 0.5 M NaOH, followed by several rapid washes in ddH_2O). The DNA was freed from the paper by incubating the paper in Hi-Salt NET Buffer (1 mM NaCl, 0.1 mM EDTA, 20 mM Tris [pH 8.0]) for 1 hour at 65 *C*, phenol extracted, ethanol precipitated, and resuspended in TE pH 7.5.

Radio-labeling DNA Fragments

Probes were prepared by either of two methods: filling in with the large fragment of *E. coli* DNA polymerase I or end labeling with polynucleotide kinase.

Table 1. Oligonucleotides

Name	Sequence
HB01	5'-ATGTGTGGAATTGTGAGCGG-3'
HB09	5'-GGCATGCCAACATGTATGGGTTCCAAGCTTG-3'
HB10	5'-CAAGCTTGGAACCCATACATGTTGGCATGCC-3'
HB39	5'-GCTAAGCTTAGCTTCCTCTATTGATGGCATGCC-3'
HB40	5'-GGCATGCCATCAATAGAGGAAGCTAAGCTTAGC-3'
HB57	5'-GACGAATTCTGCAGGGCCCGAN$_{25}$GCCAAGCTTAGCATGCACGGCC-3'
HB58	5'-CGTGCATGCTAAGCTTG-3'
HB59	5'-CGAATTCTGCAGGGC-3'
HB61	5'-GGAAGCTTGACTTTTCAGCTTCCTCTATTGATGGCATGCGGATCCGC-3'
HB62	5'-GGAAGCTTGACTTCCTGTCTTCCTATTGATTGCGCATGCGGATCCGC-3'
HB63	5'-GGAAGCTTACAATATGGACTTCCTCTTTTCTGGGCATGCGGATCCGC-3'
HB64	5'-GGAAGCTTCTAATCCGAGCTTCCACTAGGATAGGCATGCGGATCCGC-3'
HB65	5'-GGAAGCTTAGACATCGGGCTTCCACAATTTTCGGCATGCGGATCCGC-3'
HB66	5'-GGAAGCTTTTTTTCTGGCATCCAGTTTTTAATGCATGCGGATCCGC-3'
HB76	5'-GGAAGCTTCTTTTTTTACTCTTCCAGATTTTCTCGCATGCGGATCCGC-3'
HB77	5'-GGAAGCTTTCCCCTCTTTCTTCCTCTAGGGTGTGCATGCGGATCCGC-3'
HB79R	5'-GGAAGCTTTGGTGCAGGGCTTCCTCAGGTAGACGCATGCGGATCCGC-3'

Most DNA fragments used for probes in experiments were labeled by first digesting the fragment with restriction endonucleases which produce protruding 5' ends. The overhang was then filled in and labeled by using the large fragment of *E. coli* DNA polymerase I in the presence of $\alpha^{32}P$ dATP (3,000 mCi/mM). The reaction was carried out in the restriction digest reaction mixture, after digestion, by adding dNTPs (minus dATP) at 10^{-3}M. One to six units of the large fragment of *E. coli* DNA polymerase was then added and the reaction was allowed to proceed at 37 C for 30 min. Probes were then purified by 6% PAGE as described below.

In some cases probes were labeled using polynucleotide kinase. In these cases this was accomplished by first treating the DNA fragments with 0.1 U of bacterial alkaline phosphatase for 30 min at 60 C in BAP buffer (50 mM Tris [pH 8.0], 1 mM $ZnCl_2$). The reaction mixture was then treated with Proteinase K (25 µg) for 30 min at 37 C followed by two phenol extractions. The sample was then ethanol precipitated. The pellet was resuspended in polynucleotide kinase buffer (50 mM Tris [pH 7.6], 1 mM MgCl2, 5 mM DTT, 0.1 mM spermidine, 0.1 mM EDTA) with 200 µCi of $\gamma^{32}P$ ATP and 20 U of T4 polynucleotide kinase. The reaction was allowed to proceed for 1 hour at 37 *C*. Probes were then purified by 6% PAGE as described above.

Disintegrations/min were determined by suspending 1 µL in 5 ml of ddH_2O and measuring Cerenkov radiation occurring in the solution and wall of the scintillation vial (Jelley, 1958) with a scintillation counter.

Polyacrylamide Gel Electrophoresis

Various types of polyacrylamide gels electrophoresis (PAGE) were used throughout the course of this study. Nondenaturing PAGE was used for separation and purification of DNA and for band shift assays. The percentage and ratio of acrylamide to N,N'-methylenebisacrylamide (bis) was determined by the particular use of the gel. For purification of DNA, generally 6% or 8% gels with 30:1 acrylamide:bis ratio polymerized in TBE buffer (0.1 M Tris [pH 7.5], 0.1 M Boric Acid, 10 mM EDTA) were used. Running buffer was 1 x TBE. Gels were run at constant voltage, usually 1-8 V/cm.

For band shift assays, 5-10% non-denaturing gels with 82.6:1 acrylamide:bis polymerized in 0.5 x TBE buffer were commonly used. Running buffer was 0.5 x TBE. In cases were *GCR1* was translated in rabbit reticulocyte lysate, TE (10 mM Tris [pH 7.5], 1 mM EDTA) gels produced better resolution. These gels ranged between 5-10% with a acrylamide:bis ratio of 82:1 and were polymerized in TE buffer. Running buffer was TE with recirculation. Band shift gels were pre-run for at least 1.5 hours at 100 volts.

For DNA sequencing and DNase I footprinting, denaturing gels containing 7 M urea were used. An acrylamide:bis ratio of 30:1 was polymerized in 0.5 x TBE buffer. These gels were pre-run for at least 30 min at 1500 volts.

For separation of proteins 10% SDS-PAGE were used containing 0.1% SDS. The 3% stacking and 10% resolving gels contained 37.5:1 acrylamide:bis.

Determination of DNA Concentrations

DNA concentration were determined by spectrophotometry in a Beckman DU-70 spectrophotometer outfitted with a micro cell using $\varepsilon_{260} = 1.3 \times 10^4\ M^{-1}$ (per mole bp) as described in Fried and Crothers (1981).

Generation of Double Strand DNA Oligonucleotides

Double-stranded oligonucleotides were generated by three different methods. The most common protocol was by the method of Oliphant et al. (1987) as modified by Scott (1992). Single-stranded oligonucleotides were synthesized (University of Florida Interdisciplinary Center for Biotechnology Research) with the desired sequence flanked by restriction sites. The restriction site at the 3' end was contained in a larger (usually 8-10 base pair) palindrome by the addition of G and C residues. The oligonucleotides were heated in 10 µL of 3x Buffer (30 mM Tris [pH 7.5], 150 mM NaCl, 30 mM $MgCl_2$, 15 mM DTT, 0.1 mg/ml) at 37 C for 60 mins. to allow the 3' ends to self-anneal. The solution was then diluted to 30 µL with the addition of 9.5 µL of ddH_2O and 7.5 µL of dNTPs (10^{-3}M) and the 3'-ends were extended with the addition of 2 µL (10 units) of the large fragment of *E. coli* DNA polymerase I. The double-stranded extension products were then gel purified after 8% polyacrylamide (40:1.3 Acrylamide/Bis) gel electrophoresis (PAGE), digested with appropriate restriction enzymes, and cloned into the polylinker region of pUC18 or pUC19 (Messing, 1983). This method was used to generate double-stranded oligonucleotides from HB61, HB62, HB63, HB64, HB65, HB66, HB76, HB77, and HB79.

In some instances oligonucleotides were made double-stranded by directly annealing equimolar concentrations of complementary oligonucleotides. Hybridization of the oligonucleotides was determined spectrophotometrically by measuring the change in the hypochromatic shift (Bloomfield et al., 1974; Eisenberg and Crothers, 1979). Fifty micrograms of each oligonucleotides were mixed, heated to 100 C for 5 min, and then allowed to slow cool to ambient temperature. This method was used to generate double-stranded oligonucleotides from HB09/HB10, and from HB39/HB40 (see Baker, 1991).

Finally, some oligonucleotides were made double-stranded by annealing a smaller primer to a longer template oligonucleotide and extending with dNTPs (10^{-3} M) in the presence of the large fragment of E. coli DNA polymerase I at 37 C for 30 min using 3x Buffer (described above). This method allowed incorporation of the polymerase chain reaction (PCR) for re-amplification of the template oligonucleotide by using an additional smaller primer. PCR cycle temperatures and times were as follows: denaturation, 94 C for 50 sec; annealing, 50 C for 40 sec; and extension, 72 C for 35 sec. This method was used to generate double-stranded oligonucleotides from HB57, HB58, and HB59.

DNA Sequencing

DNA sequencing was carried out by the dideoxy chain termination method of Sanger (1977) as modified by U.S. Biochemicals (Cleveland, Ohio) utilizing the Sequenase enzyme, with the following addition: After plasmids were

prepared from 10 ml cultures of *E. coli* by the alkaline lysis method (Birnboim and Doley, 1979; Ish-Horowwicz and Burke, 1981) they were further purified by resuspending the total precipitated DNA from the preparation in 128 µL ddH_2O. Then 32 µL of 4 M NaCl and 160 µL of 13% polyethylene glycol-8000 (PEG) were added and the solution was mixed well. The solution was incubated on ice for exactly 20 mins. and then microcentrifuged for 10 mins. at room temperature. The pellet was washed with 70% ethanol and then dried under vacuum. The pellet was resuspended in 75 µL of TE pH 7.5. 25 µL of this solution was then used in the sequencing method described by U.S. Biochemicals.

Plasmid Construction

The plasmids used in this study are listed in Table 2. A number of plasmids which contain CTTCC sequence elements were constructed. Plasmid pUC66 contains sequence from UAS_{TPI}, from positions -317 to -327. This region harbors two GCR1-binding sites and a RAP1-binding site (Scott, 1992). Plasmid pUC66 was constructed by cloning the *Hind*III-*Sph*I fragment from plasmid pES119 (Scott, 1992) into the *Hind*III-*Sph*I sites of pUC18.

Plasmids pUCT61-pUCT79 contain CTTCC sequence elements from a number of glycolytic and translational machinery genes. The number of the plasmid refers to the HB oligonucleotide from which the insert was derived (sequences are listed in Table 1). After the oligonucleotides were made double-stranded (described above), 5 µg of DNA was digested with *Hind*III and *Bam*HI and cloned into the *Hind*III-*Bam*HI

Table 2. Plasmids

Plasmid	Comments
pHB66 (Baker, 1991)	*GCR1* structural gene cloned downstream of SP6 promoter in plasmid pSP19
pSP56RT (Chambers et al, 1987)	*RAP1* structual gene cloned downstream of SP6 promoter in plasmid pSP19
pMH2	*Pst*I-*Sal*I fragment of *GCR1* structural gene cloned into *Pst*I-*Sal*I sites of pSP18
pMAL-GCR1(690-844)	*malE-GCR1*(690-844) fusion gene under *tac* promoter in pMAL-c
pMAL-GCR1(783-844)	*malE-GCR1*(783-844) fusion gene under *tac* promoter in pMAL-c
pCD1	*malE-GCR1*(1-844) fusion gene under *tac* promoter in pMAL-c
pCD2	*malE-GCR1*(277-844) fusion gene under *tac* promoter in pMAL-c
pCD3	*malE-GCR1*(422-844) fusion gene under *tac* promoter in pMAL-c
pCD5	*malE-GCR1*(706-844) fusion gene under *tac* promoter in pMAL-c
pUCT61	CTTCC sequence element from *TPI1* cloned into *Hind*III-*Bam*HI site of pUC18
pUCT62	CTTCC sequence element from *PGK1* cloned into *Hind*III-*Bam*HI site of pUC18
pUCT63	CTTCC sequence element from *ADH1* cloned into *Hind*III-*Bam*HI site of pUC18
pUCT64	CTTCC sequence element from *ENO1* cloned into *Hind*III-*Bam*HI site of pUC18
pUCT65	CTTCC sequence element from *PYK1* cloned into *Hind*III-*Bam*HI site of pUC18

Table 2. Plasmids (cont.)

Plasmid	Comments
pUCT66	CATCC sequence element from *TPI1* cloned into *Hind*III-*Bam*HI site of pUC18
pUCT76	CTTCC sequence element from *TEF1* cloned into *Hind*III-*Bam*HI site of pUC18
pUCT77	CTTCC sequence element from *TEF1* cloned into *Hind*III-*Bam*HI site of pUC18
pUCT79	CTTCC sequence element from *CRY1* cloned into *Hind*III-*Bam*HI site of pUC18
pUC66	UAS_{TPI1} cloned into *Hind*III-*Sph*I site of pUC18
pCY4 (Prentki et al., 1987)	Circular permutation assay vector
pCY66	UAS_{TPI1} cloned into *Sma*I-*Bgl*II site of pCY4

Table 3. Oligonucleotides containing CTTCC sequences

Oligonucleotide	Sequence
LINKER	agcttgcatgcctgcaggtcgactctagaggatccccgggtaccgagctcgatt
UASTPI	agcttAGCTTCCTCTATTGATGTTACACCTGGACACCCCTTTTCTGGCATCCAGTTgcat
$TPI1_1$	agcttGACTTTTCAGCTTCCTCTATTGATGgcatgcggatccccgggtaccgagctcgatt
$TPI1_2$	agcttt-TTTTCTGGCATCCAGTTTTTAATgcatgcggatccccgggtaccgagctcgatt
PGK	agcttGACTTCCTGTCTTCCTATTGATTGCgcatgcggatccccgggtaccgagctcgatt
ENO1	agcttCTAATCCGAGCTTCCACTAGGATAGgcatgcggatccccgggtaccgagctcgatt
PYK	agcttAGACATCGGGCTTCCACAATTTTCGgcatgcggatccccgggtaccgagctcgatt
ADH1	agcttACAATATGGACTTCCTCTTTTCTGGgcatgcggatccccgggtaccgagctcgatt

sites of pUC18. The inserts of these plasmids, used for radiolabeled DNA probes in bands shift assays, are listed in Table 3.

A series of *malE::GCR1* gene fusions which carry various deletions of the 5' end of the *GCR1* structural gene were prepared by cloning Gcr1p coding sequences into the plasmid pMAL-c (Guan et al., 1987; Maina et al., 1988). The plasmid pMAL-c express the *malE* gene, which encodes the *E. coli* maltose binding protein (MBP), under control of the *E. coli* tac promoter (Amann et al., 1983). A polylinker site is located in the *malE* structural gene and allows in-frame insertion of DNA fragments to construct fusion proteins under a strong inducible promoter. Plasmid pMAL-GCR1(690-844), which encodes a maltose binding protein (MBP)-Gcr1p fusion protein containing amino acids 690-844 of Gcr1p, was created by cloning the *Hae*III-*Xba*I fragment from plasmid pMH2 into the *Stu*I-*Xba*I sites of plasmid pMAL-c. Plasmid pMH2 contains the *Pst*I-*Sal*I genomic fragment of the *GCR1* structural gene cloned into the *Pst*I-*Sal*I sites of plasmid pSP18.

Plasmid pMAL-GCR1(783-844) was created as follows: Plasmid pMAL-GCR1(690-844) was cleaved with *Ppmu*I. The resulting overhang was filled in with dNTPs and the large fragment of *E. coli* polymerase I. After phenol extraction and ethanol precipitation the DNA was then cleaved with *Hind*III. The *Ppmu*I-*Hind*III fragment was isolated by 8% PAGE and then cloned in-frame into the *Stu*I-*Hind*III sites of pMAL-c.

Plasmids pCD1, pCD2, pCD3, and pCD5 were kindly provided by Carolyn M. Drazinic. These plasmids code for additional 5'-GCR1 deletions fused in frame with MBP. They contain the following amino acids residues of GCR1: pCD1, MBP-GCR1(1-844); pCD2, MBP-GCR1(277-844); pCD3, MBP-GCR1(422-844); pCD5, MBP-GCR1(706-844).

In Vitro Transcription

Plasmid pHB66 (Baker, 1991) contains the GCR1 structural gene (from an *Afl*II restriction site located 136 bp 5' to the translational start site to a *Bcl*I site located 661 bp 3' to the translational termination site) cloned downstream of the SP6 promoter in the plasmid pSP19. This plasmid was linearized with various restriction enzymes. After complete digestion was confirmed by agarose gel electrophoresis, the DNA was phenol extracted, ethanol precipitated, and resuspended in TE. For a translation template for RAP1 the plasmid pSP56RT (Chambers et al., 1989) was linearized by cleavages with the restriction endonuclease *Xba*I, and was prepared in a similar manner.

Five micrograms of linearized DNA was used as a template in transcription reactions carried out in the presence of the cap analog $m^7G(5')ppp(5')G$ using a kit from Promega, under the following reaction conditions: 40 mM Tris [pH 7.5], 6 mM $MgCl_2$, 2 mM spermidine, 10 mM NaCl, 10 mM DTT, 0.1 mg/ml BSA, 1 mM ATP, CTP, UTP, 0.1 mM GTP, 0.5 mM GpppG and 50 U or RNasin. After transcription reactions were incubated with SP6 RNA polymerase at 37 C for 60 min, the samples were treated with DNase I for 15 min., phenol/chloroform

extracted, ethanol precipitated, and then resuspended in 20 µL of ddH2O and stored at -70 *C*.

In vitro Translation

In vitro-derived transcripts were translate in a rabbit reticulocyte lysate system in the presence of L-[^{35}S]methionine by using a kit obtained from Promega. Two microliters of substrate RNA (prepared as described above) was incubated with 18 µL of nuclease treated rabbit reticulocyte lysate, 4 µL ddH2O, 1 µL RNasin (50 U), 0.5 µL 1 mM amino acid mixture (minus methionine), and 2.0 µL of [^{35}S]methionine. The amount of Gcr1p produced in the rabbit reticulocyte lysates was estimated by determining the amount of [^{35}S]-methionine incorporated into trichloroacetic acid-precipitable material. One microliter of the lysate was incubated with 50 µL of 0.1 N NaOH at 37 *C* for 15 min and then added to 1 ml of 10% TCA and placed on ice for 30 min. Samples were then vacuum filter through filter paper, and washed with 1 ml of 10% TCA. Filters were then added to 5 ml of scintillation fluid and counted. Typical *in vitro* translation reactions yielded approximately 2.4 ng of GCR1 per µL of rabbit reticulocyte lysate. The translation products were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), and the radiolabeled proteins were visualized by autoradiography at -70 *C* and compared with known molecular weight standards.

In Vitro DNase I Protection Assays

GCR1 binding sites in the UAS_{TPI} were determined by DNase I footprints, carried out by a modification of the

method of Pfeifer et al. (1987). A 234 bp DNA fragment from the UAS_{TPI} was generated by cleavage of the plasmid pES37 (Scott, 1992) with the restriction endonucleases *Hind*III and *Sph*I. This fragment was then radiolabeled only at the *Hind*III end by filling in with the large fragment of *E.coli* DNA polymerase I in the presence of α^{32}P dATP. The DNA probe was isolated from the plasmid by 8% PAGE. After visualization of the ethidium bromide stained gel by long wave UV, the area of the gel containing the DNA probe was excised, and the DNA extracted from the gel by the DEAE paper method as described above.

Nucleoprotein complexes were allowed to form under standard band-shift reaction conditions in a total volume of 20 µL, and then various amounts of DNase I (0.2 to 0.5 U) were added. The reaction mixture was incubated at ambient temperature for 2 min. The reaction was stopped by the addition of an equal volume of stop solution (50% glycerol, 0.25 M EDTA [pH 8.0]). The resulting mixture was immediately loaded onto a native 5% polyacrylamide (49.4:0.6) 0.5 x TBE gel. Free and complexed DNA were detected by autoradiography of the wet gel. Bound DNA from both the upper and lower complexes was cut out of the gel and eluted from the gel by either soaking overnight at 37 C in 500 mM ammonium acetate-1 mM EDTA or by the DEAE paper method described above. The eluted DNA was then precipitated, washed with 70% ethanol, suspended in 10 µL of loading buffer (95% Formamide, 20 mM EDTA, 0.05% Bromophenol Blue, and 0.05% Xylene Cyanol FF), boiled for 5 min, and loaded onto a denaturing sequencing

gel. Depending on the experiment 10-20,000 counts/min were loaded per lane.

DNA Band-Shift Assays

Protein-DNA complexes were studied using modifications of the procedures of Fried and Crothers (1981) and Garner and Revzin (1981). The binding reactions were incubated for 20 minutes in a 20 µL volume at ambient temperature in binding buffer composed of 12 mM HEPES (pH 7.5), 60 mM KCl, 5 mM $MgCl_2$, 4 mM Tris, 0.6 mM EDTA, 0.6 mM DTT, 10% glycerol, and 0.3 µg/µL BSA. In some cases poly (dI/dC) was used as a non-specific competitor at 0.26 µg/µL. In some case double-stranded oligonucleotides were used as competitor. Competitors were diluted and added as a constant volume per reaction mixture. Protein extracts used in experiments were diluted with binding buffer. Components were added in the following order: binding buffer, competitor, probe, protein. In qualitative experiments, the amount of DNA added was based upon disintegrations/min, usually between 10-20,000 counts/min. Depending upon the experiment, gels varied between 5-10% polyacrylamide. Buffers were either TE (0.1 M Tris [pH 7.5], 5 mM EDTA) or 0.5 TBE (0.05 M Tris [pH 7.5], 0.05 M Boric Acid, 5 mM EDTA). In the case of TE the running buffer (10 mM Tris [pH 7.5], 1 mM EDTA) was recirculated. All gels were pre-run for at least 1.5 hours at 100 volts. Samples were applied to the running gel. After all samples were loaded the voltage was increased to 150 volts and the gels were run until the adequate separation desired (based on pilot experiments).

Titration of DNA-Binding Activity

DNA-binding activity of MBP-Gcr1p(690-844) was determined by band shift assay using the approach of Riggs et al. (1970) and Chadwick et al. (1970). A constant amount of purified MBP-Gcr1p(690-844) fusion protein was titrated with increasing amounts of radiolabeled DNA containing a Gcr1p-binding site. Free and bound complexes were then separated by a standard band shift assay. The amount of DNA used was accurately determined by spectrophotometry. The amount of DNA that was shifted was determined by using phosphoImager analysis (Molecular Dynamics). Thus, the amount of DNA in the shifted complex was determined by comparison to a standard curve of known concentrations of DNA. From this analysis the amount of DNA retained at plateau is a measure of total concentration of active protein in terms of DNA concentration.

Calculation of Equilibrium Binding Constant

DNA-binding affinity was determined by band shift assay using the analysis of Riggs et al. (1970) and Chadwick et al. (1970). A known amount of DNA was radiolabeled and allowed to bind to varying concentrations of active MBP-Gcr1p. Free DNA was separated from nucleoprotein complexes by native gel electrophoresis. Gels were dried to Whatman paper and then counts were determined by use of a PhosphoImager (Molecular Dynamics). By holding DNA concentrations below the estimated K_d (determined from pilot experiments), the apparent K_d could be determined by the amount of protein added when the counts of free DNA equalled the counts in the nucleoprotein complex.

RESULTS

> The world is a construct of our sensations, perceptions, memories. It is convenient to regard it as existing objectively on its own. But it certainly does not become manifest by its mere existence.
>
> Erwin Schrodinger (p.1)

Previous experiments had suggested that the product of *GCR1* interacts directly with the CTTCC sequence element (Baker, 1991). In those experiments, RNA was translated *in vitro* to make the Gcr1p protein in rabbit reticulocyte lysates. The lysate was used in band shift assays with DNA fragments containing the CTTCC sequence element. Additionally, when anti-Gcr1p antibody was included in the reaction mixture a supershift was observed (Baker, 1991). These experiments strongly supported the view that Gcr1p is a DNA-binding protein; however, alternative models could not totally be excluded. For example, since CTTCC sequence elements had been described in the promoters of higher eukaroytes it was possible that the Gcr1p could associated with and modify another factor in the lysate stimulating a protein complex which included Gcr1p to bind to UAS_{TPI} DNA. For example, the product of the retinoblastoma gene displays DNA-binding activity when complexed to the factor EF2 (Chittenden et al., 1993; Ray et al., 1992; Ouellette et al., 1992). Alternatively, it could be argued that the *GCR1* gene product allosterically modified another protein unleashing

binding activity of the complex. If this were the case, then Gcr1p would be included in the complex--based on the supershift experiments. This study set out to confirm that Gcr1p interacts directly with DNA in a sequence-specific manner.

Gcr1p Expressed in *E. coil* or Rabbit Reticulocyte Lysate Binds to DNA

Throughout the course of this study Gcr1p, synthesized *in vitro* from rabbit reticulocyte lysates and MBP-Gcr1p fusion protein, expressed in *E. coli*, were used to characterize the DNA binding activity of Gcr1p.

MBP-Gcr1p full-length fusion protein was produced in *E. coli* strain TB1 harboring a plasmid encoding for a *malE::GCR1*(1-844) gene fusions under the *tac* promoter. The fusion protein was induced in *E. coli* by the addition of 2 mM IPTG during the log phase of growth. Cell were then lysed by passage through a French Pressure cell. Protein extracts were analyzed by SDS-PAGE stained with Coomassie blue (data not shown). Protein concentration, pre-determined by the method of Bradford (1976), ranged from 0.068 to 0.2 mg/ml. The amount of extract used in band shift assays varied depending on the preparation. Typically between 3 and 15 microliters were used in band shift assays. The volume used was determined by titrating DNA-binding activity. It was noted that if extracts were used immediately less volume was necessary to detect binding. If extract was used after overnight storage at -20 C then a larger volume was necessary for detectable activity (data not shown). This was

presumably due to instability of full-length Gcr1p in the extracts.

The DNA probe used for the band-shift assay was the upstream activating sequence (UAS) of the gene encoding triose-phosphate isomerase (*TPI*1). Plasmid pES119 (Scott, 1992) containing the UAS_{TPI1}, was digested with the restriction endonucleases *Sph*I and *Hind*III liberating a 60 base pair fragment containing UAS_{TPI1} from positions -377 to -327 relative to the start of translation. The fragment was then radiolabeled at the *Hind*III site.

As seen in Figure 1, both Gcr1p synthesized *in vitro* and the full length MBP-Gcr1p fusion protein were able to form nucleoprotein complexes with DNA carrying UAS_{TPI1}. The position of the nucleoprotein complexes observed with the fusion protein migrate more slowly than the complexes observed with rabbit reticulocyte lysate containing Gcr1p (Figure 1). This difference is presumably due to the increase in molecular weight of the nucleoprotein complex due to the presence of the maltose binding moiety of the fusion protein. This is consistent with the notion that Gcr1p interacts directly with UAS_{TPI1}. The appearance of two shifted bands in the band shift assay with Gcr1p and its derivatives was routinely observed and had been noted previously (Baker, 1991). It has also recently been shown that Gcr1p synthesized in yeast also results in two shifted complexes in the band shift assay (Willett et al., 1993). The reason for two bands is currently unknown. Figure 1 also shows that no nucleoprotein complexes were observed with

Figure 1. Comparison of DNA-binding activity of Gcr1p and hybrid MBP-Gcr1p fusion protein. DNA band-shift assays were carried out with a radiolabeled fragment of DNA carrying UAS_{TPI1} (see Table 3. for sequence). The radiolabeled DNA was incubated in binding buffer with protein extract indicated above the lanes as described in Material and Methods. Fragment alone, radiolabeled DNA fragment; No RNA RRL, 5 µl of untreated rabbit reticulocyte lysate (RRL); Gcr1p RRL, 5 µl of rabbit reticulocyte lysate containing *in vitro* synthesized Gcr1p; *E. coli*/pMAL-cR1 Uninduced, 1 µl of an extract of a *E. coli* culture harboring plasmid pMAL-cR1; *E. coli*/pMAL-cR1 Induced, 1 µl of an extract of a *E. coli* culture harboring plasmid pMAL-cR1 which had been induced with IPTG 2 h prior to harvest; *E. coli*/pMAL-Gcr1p(1-844) Uninduced, 1 µl of an extract of a *E. coli* culture harboring plasmid pMAL-Gcr1p(1-844) (numbers in parentheses denote amino acid residues of Gcr1p present in the expressed polypeptide, see Table 2.); *E. coli*/pMAL-Gcr1p(1-844) Induced, 1 µl of an extract of a *E. coli* culture harboring plasmid pMAL-Gcr1p(1-844) which had been induced with IPTG 2 h prior to harvest; f, free unbound probe.

Frag. alone

NO RNA RRL

Gcr1p RRL

E.coli/pMAL-cR1 Uninduced

E.coli/pMAL-cR1 Induced

E.coli/pMAL-Gcr1p(1-844) Uninduced

E. coli/pMAL-Gcr1p(1-844) Induced

untranslated rabbit reticulocyte lysates, *E. coli* extracts prepared from uninduced and induced cultures of strains carrying the parent plasmid, pMAL-cR1, and extracts prepared from uninduced cultures of *E. coli* strains carrying the *malE::GCR1* gene fusion.

The DNA-Binding Domain of Gcr1p Resides within the Carboxy-Terminal 154 Amino Acid Residues

To map the DNA-binding domain of Gcr1p a series of DNA band-shift experiments with truncated versions of Gcr1p were performed. Carboxy-terminal truncations of Gcr1p were synthesized *in vitro*. Plasmid pHB66 was linearized with a series of restriction endonucleases that cleaved within the *GCR1* structural gene as shown in Figure 2. The linearized constructs were then used as templates for *in vitro* RNA synthesis, and run off transcripts were translated *in vitro* using rabbit reticulocyte lysates. Production of polypeptides of the desired molecular mass was confirmed by SDS-PAGE and autoradiography (Figure 3). The rabbit reticulocyte lysates were then used in a series of band-shift experiments to determine which, if any, of the truncated fusion-proteins had DNA-binding activity. Figure 4 shows that only full length Gcr1p had binding activity. This suggested that the DNA-binding domain resided in the carboxy terminus of the polypeptide.

To test this possibility, a similar set of experiments with an amino-terminal deletion series was carried out. For these experiments, plasmids pCD1, pCD2, pCD3, and pCD5 (coding for various gene fusions--see Material and Methods)

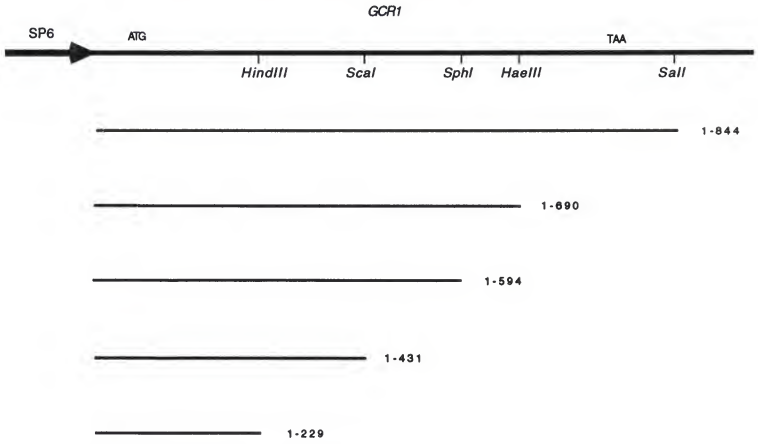
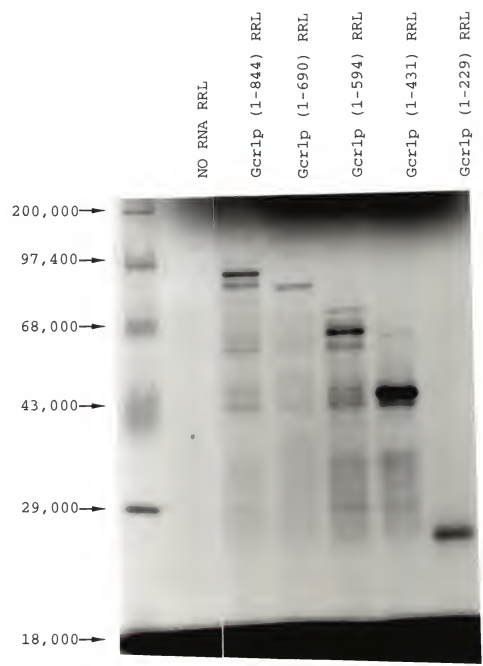


Figure 2. Templates used to generate carboxy-terminal deletions of Gcr1p. A schematic representation of the GCR1 structuaral gene cloned downstream of the SP6 promoter is displayed. Digestion of the construct with the various enzymes shown produce RNA templates coding for polypeptides of the length display below (numbers refer to amino acids residues of Gcr1p).

Figure 3. Carboxy-terminal truncation polypeptides of Gcr1p. Autoradiography of a 10% SDS-polyacrylamide gel electrophoresis of polypeptides produced by *in vitro* translation of the RNA templates (displayed in Figure 2) using rabbit reticulocyte lysate (RRL). Polypeptides were translated *in vitro* in the presence of ^{35}S-methionine. The numbers correspond to amino acids residues of Gcr1p. Molecular weight standards, in kiloDaltons, are as follows: myosin H-chain, 200,000; phosphorylase b, 97,000; bovine serum albumen 68,000; ovalbumin, 43,000; carbonic anhydrase, 29,000; β-lactoglobulin, 18,000.


NO RNA RRL
Gcr1p (1-844) RRL
Gcr1p (1-690) RRL
Gcr1p (1-594) RRL
Gcr1p (1-431) RRL
Gcr1p (1-229) RRL
200,000
97,400
68,000
43,000
29,000
18,000

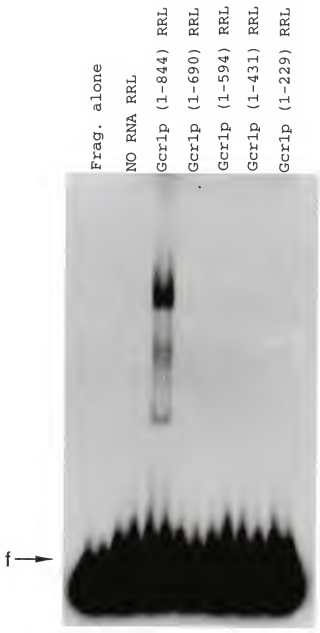


Figure 4. DNA-binding activity of Gcr1p carboxy-terminal deletion polypeptides. A typical DNA band-shift assay using the *in vitro* synthesized Gcr1p carboxy-terminal deletion series is shown. The numbers in parentheses denote the amino acid residues of Gcr1p expressed in the rabbit reticulocyte lysates (RRL). Five microliters of each rabbit reticulocyte lysate was used in the assays.

were provided by Carolyn Drazinic. These plasmids contain subgenic portions of *GCR1* fused in-frame to *malE* in the fusion vector pMAL-cR1 (see Figure 5). *E. coli* lysates from induced strains harboring the *malE::GCR1* gene fusions were prepared. Production of MBP-Gcr1p hybrid polypeptide of the desired molecular mass was confirmed by SDS-PAGE and the fusion-protein visualized by coomassie blue staining (data not shown). In general, there was an inverse correlation between the size of the fusion polypeptide and the amount of material observed. *E. coli* lysates containing hybrid MBP-Gcr1p polypeptide were then tested in DNA-band-shift assays (Figure 6). All constructs that carried the carboxy-terminal 154 amino acid residues of Gcr1p were able to bind to the upstream activating sequence of *TPI1*. Whereas lysates from strains expressing fusion protein carrying the carboxy-terminal 138 amino acids of Gcr1p were unable to bind the DNA fragment used in the study.

A summary of the mapping data is shown in Figure 7. From these results it was concluded that the DNA-binding domain of Gcr1p resides somewhere within the carboxy-terminal 154 amino acids of Gcr1p.

DNase I Footprint Analysis of UAS_{TPI}

To establish that Gcr1p indeed bound to the CTTCC sequence element, *in vitro* DNase I footprinting experiments were performed. For these experiments purified preparations of the smallest fusion protein which retained DNA-binding activity, MBP-Gcr1p(690-844), were used. The smaller fusion proteins were more stable and could be produced in larger

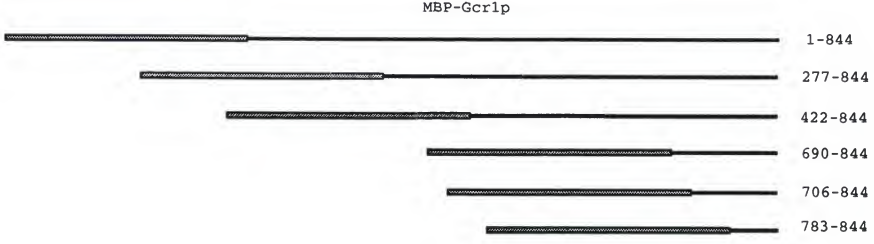
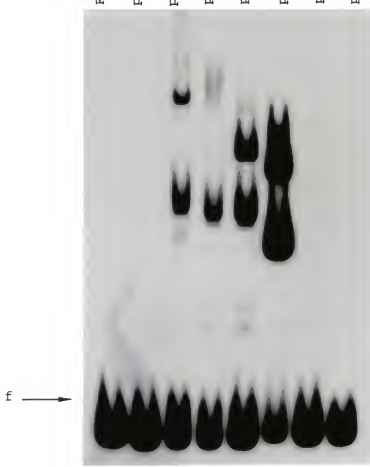


Figure 5. *malE::GCR1* gene fusions products. A schematic representation of the Gcr1p polypeptides used in the amino-terminal deletion studies. The stippled lines indicate maltose-binding protein moiety carried in the fusion proteins. Numbers indicate amino acid residues of Gcr1p conatained in the fusion proteins

Figure 6. DNA-binding activity of MBP-Gcr1p hybrid proteins. A DNA band-shift assay using a Gcr1p amino-terminal deletion series of hybrid MBP-Gcr1p fusion proteins is displayed. One microliter of an extract of the induced *E. coli* culture, indicated above each lane, was added to the standard band-shift reaction mixture, as described in the legend to Figure 1.


Frag. alone
pMAL cRl
pMAL-Gcr1p(1-844)
pMAL-Gcr1p(277-844)
pMAL-Gcr1p(422-844)
pMAL-Gcr1p(690-844)
pMAL-Gcr1p(706-844)
pMAL-Gcr1p(783-844)
f

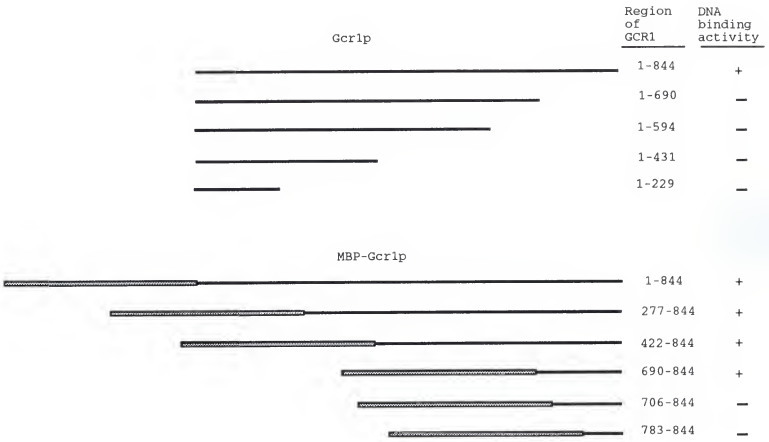


Figure 7. Summary of data mapping the Gcr1p DNA-binding domain. The figure is a schematic representation of the Gcr1p polypeptides used in the mapping study and summary of results obtained. Solid lines represent regions of Gcr1p carried in the polypeptide. The stippled lines indicate maltose-binding protein moiety carried in the fusion protein.

quantity. The stability of the MBP-Gcr1p(690-844) protein also allowed for its purification. The fusion protein was purified as described in Material and Methods. A 234 base-pair DNA fragment from the UAS_{TPI1} was generated by cleavage of the plasmid pES37 (Scott, 1992) with the restriction endonucleases *Sph*I and *Hind*III. The fragment was radiolabeled at the *Hind*III site. Nucleoprotein complexes were allowed to form under standard band shift reaction conditions after which DNase I was added to the mixture. Reactions were terminated by inhibiting DNase I with high concentrations of EDTA, and then the samples were immediately run into a non-denaturing acrylamide gel. Complexes were identified by autoradiography and purified from the gel. The nicked DNA was then transferred to DEAE paper using electrophoresis. DNA was eluted with high-salt buffer and then isolated by phenol extraction and ethanol precipitation. The DNA was then denatured and run on a sequencing gel. A DNA fragment incubated without protein was digested with DNase I to generate a control ladder. Figure 8 shows the results of the DNase I protection studies using purified MBP-Gcr1p(690-844) fusion-protein. Two areas of protection were observed: one clear area of protection centered over the CTTCC motif and another area of partial protection centered over the related sequence CATCC. Edward Scott has shown that both of these sequences are protected *in vivo* in a *GCR1*-dependent manner by *in vivo* G methylation protection experiments (Huie et al., 1992; Scott, 1992; Scott and Baker, 1993).

Figure 8. Gcr1p DNA-binding domain protects the CTTCC sequence motif and a related sequence element, CATCC, in UAS_{TPI1} from DNase I cleavage. Analysis of UAS_{TPI1} was carried out with purified MBP-Gcr1p(690-844) and a radiolabeled 234-bp fragment carrying the UAS of *TPI1*. Lanes T, G, C, and A, are the products of the dideoxy sequencing reactions of M13mp18 and serve as molecular weight standards; lanes 1 and 4, free fragment treated with DNase I; lane 2, nucleoprotein complex treated with 0.2 U of DNase I, as described in Material and Methods; lane 3, nucleoprotein complex treated with 0.5 U DN ase I. The sequences of protected areas are denoted on the right. The exact extent of the area protected over the CTTCC sequence element could not be determined because of lack of bands in the control lanes (lanes 1 and 4); therefore, two 5' boundaries are indicated on the figure.

T G C A 1 2 3 4
5'-TAAGACTTTTCAGCTTCCTCTAT-3'
5'-TTCTGGCATCCAGTTT-3'

Identification of a Consensus Gcr1p DNA Binding Sequence

Before this study, demonstration of the DNA-binding activity of Gcr1p had utilized only DNA carrying sequence from the *TPI1* UAS element. In an effort to define a consensus Gcr1p-binding site, DNA was synthesized with the putative Gcr1p-binding sites found in front of a number of other genes encoding glycolytic enzymes, namely *PGK1*, *ENO1*, *PYK1*, and *ADH1*. These sequences were chosen for the following reasons. There are three CTTCC sequence elements important for expression of *PGK1* (Chambers et al., 1988). One of these elements was arbitrarily chosen for study. The CTTCC elements from *ENO1* and *PYK1* were chosen because there is genetic evidence that shows they are important determinants of their respective UAS elements (Buchman et al., 1988). It was predicted that the CTTCC element found adjacent to the RAP1-binding site in the *ADH1* promoter would likely bind Gcr1p; therefore it was tested it for its ability to bind Gcr1p. In addition, the two putative Gcr1p-binding sites in UAS_{TPI} (determined from the DNase I footprinting experiments) were tested individually.

Oligonucleotides were synthesized which carried 25 nucleotides from the genes of interest with the CTTCC motif under study located at the center of each oligonucleotide. Table 3 shows the sequence of the oligonucleotides used as probes in this study. Due to the relative positions of the elements in UAS_{PGK}, the oligonucleotide carrying the *PGK* sequence contained two CTTCC sequence elements. The CTTCC sequence that is displayed in enlarged type from the *PGK1*

sequence listed in Table 3 has recently been shown to be protected in wild-type strains and deprotected in a *gcr1* null mutant by *in vivo* footprinting (Stanway et al.,1994). The oligonucleotides were then tested for their ability to interact with the DNA-binding domain of Gcr1p in a series of DNA-band-shift assays. As a negative control the polylinker from pUC18 was used in the band shift assays. The results of the band shift assays are shown in Figure 9. The appearance of shifted bands were not detected when the polylinker from pUC18 was used as probe. On the other hand, each of the putative Gcr1p-binding sites gave rise to the appearance of shifted bands, thus providing evidence that Gcr1p is capable of interacting with these sequences. Additionally, the UAS of *CYC1*, which does not contain a CTTCC sequence element and is not under control of *GCR1*, also did not form a nucleo-protein complex with Gcr1p (data not shown).

The expression of the genes encoding elongation factor EF-1α (*TEF1* and *TEF2*) and ribosomal protein 59 (*CRY1*) is reduced two- to four-fold in *gcr1* mutant strains (Santangelo and Tornow, 1990). Potential Gcr1p-binding sites were identified located in the 5' noncoding region of these genes. A model was entertained that Gcr1p played a direct role in expression of these genes at the transcriptional level. Therefore, oligonucleotides which carried 25 nucleotides containing the CTTCC motif from the genes of interest were synthesized. These oligonucleotides were also tested by DNA-band-shift assay for their ability to interact with the DNA-binding domain of Gcr1p. None of the oligonucleotides bound

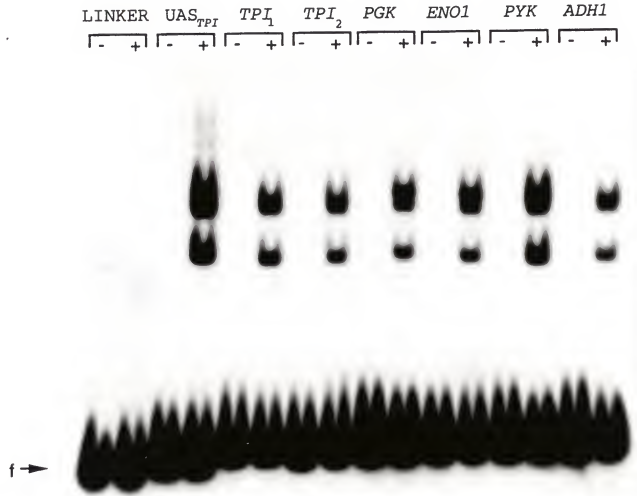


Figure 9. Gcr1p DNA-binding domain binds to CTTCC sequence elements found in front of other glycolytic genes. DNA band-shift analysis was carried out with purified MBP-Gcr1p (690-844) and the radiolabeled oligonucleotides listed in Table 3. The absence (-) and presence (+) on 10 ng of MBP-Gcr1p fusion protein in the DNA band-shift assay are indicated above each lane.

DNA with greater affinity than the polylinker of pUC18 (data not shown).

From the results of the band-shift experiments the following consensus DNA-binding sequence for Gcr1p was derived (see Discussion):

(T/A)N(T/C)N(G/A)N **C (T/A) T C C** (T/A)N(T/A)(T/A)(T/G)

The Gcr1p-DNA Nucleoprotein Complex

Having defined the DNA-binding domain of Gcr1p, and a consensus DNA-binding site which it recognizes, I set out to characterize the affinity and specificity of the Gcr1p interaction with its binding site. Historically, it had been very difficult to demonstrate that Gcr1p was a DNA-binding protein. During the period between sequencing of the gene (Baker, 1986) and the demonstration of the binding activity of it product (Baker, 1991) many laboratories had attempted to demonstrate Gcr1p DNA-binding activity, yet failed. It was therefore reasoned that Gcr1p may have a low affinity for its DNA site, and that it may be facilitated to bind DNA by other proteins. To test this I measured the binding affinity of Gcr1p for its target site.

The DNA-binding affinity of the Gcr1p-binding domain

Experience working with the Gcr1p protein revealed that the full-length protein is very unstable. This seems to be the case whether the product is translated *in vitro* in rabbit reticulocyte lysate or whether it is expressed in *E. coli*. Full-length *in vitro* translated product must be used immediately for detectable activity. Similarly, the full length MBP-Gcr1p(1-844) fusion-protein must be used

immediately. This rapid decrease in activity made it extremely difficult to use full-length Gcr1p in experiments were the active protein concentration needs to be accurately known, or in experiments in which the protein must be of a fixed concentration over time. However, truncated MBP-Gcr1p fusion-proteins including the Gcr1p DNA-binding domain has proven to be very stable, especially when stored at high concentration. I do not detect a noticeable loss of activity over months of storage at -70 C. This allowed me to measure the binding affinity of the DNA-binding domain. Fusion-proteins have been used to measure the DNA binding affinity of other proteins (Desplan et al., 1985; Johnson and Herskorwitz, 1985; Giese et al., 1991).

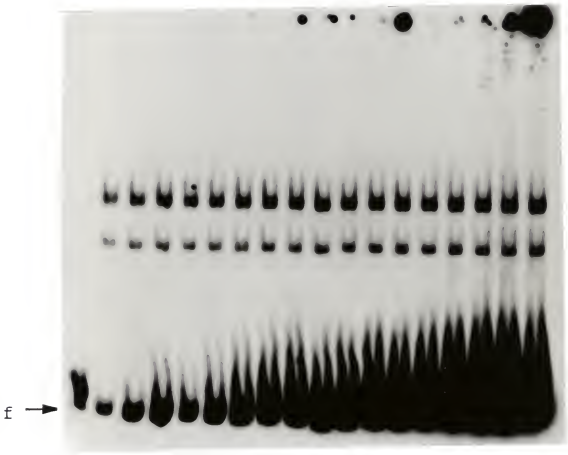
A single Gcr1p-binding site from the UAS of the triose-phosphate isomerase gene was used in these experiments. The site is listed as TPI_1 in Table 3, and, as mentioned above, it has been shown to be protected in a *GCR1*-dependent manner *in vivo* (Huie et al., 1992; Scott, 1992). The concentration of the radiolabeled, gel-purified probe was determined by spectrophotometry.

Purified hybrid MBP-Gcr1p(690-844) fusion protein was used for these experiments. Active protein concentration was determined by the methods described by Riggs et al. (1970) and Chadwick et al. (1970), except that band-shift assays were used instead of DNA filter-binding assays. The protein concentration was held constant, using 1 µL of a 1:1000 dilution of the stock solution, and increasing amounts of radiolabeled DNA was added. The concentration of the stock

solution of protein was 0.35 µg/µL as determined by the method of Bradford (1976). After complexes were allowed to form, the reaction mixtures were run into a 0.5 x TBE 5% native PAGE. The gels were dried on the Whatman paper, and then the separated radiolabeled DNA was quantitated with a PhosphoImager (Molecular Dynamics). The fraction of DNA shifted by protein was determined. At saturating conditions, all of the active protein was bound to the DNA. Since the concentration and specific activity of the DNA in the complex was known, the concentration of the complex was determined by quantitating the amount of DNA complexed. The active protein concentration therefore was expressed in DNA equivalents per volume (see Chadwick et al., 1970). A typical experiment is shown in Figure 10. In two independent experiments the active protein concentration of the stock was determined to be 1×10^{-6} M. The active fraction was approximated to be 16% of the total protein concentration.

Pilot experiments were performed to estimate the dissociation constant of DNA-binding domain of Gcr1p from its recognition site. With the protein activity known, an arbitrary small concentration of DNA was held constant at 1.2×10^{-10} M and increasing concentrations of the protein were added. The concentration at which one-half of the DNA was bound was used to estimate the binding affinity. The dissociation constant was estimated to be in the 10^{-10} M range. Therefore, the DNA concentration was maintained below this approximated dissociation constant. DNA concentrations of 6.3×10^{-12} M, 1.9×10^{-11} M, and 5.5×10^{-11} M were used in

Figure 10. Titration of Gcr1p DNA-binding activity. The concentration of active MBP-Gcr1p(690-844) fusion-protein was determined by titrating increasing concentrations of radio-labeled oligonucleotide carrying a single Gcr1p-binding site (probe TPI_1, listed in Table 3.) with a constant amount of MBP-Gcr1p(690-844). A typical experiment is shown here. As the amount of DNA was increased all of the active protein appeared in the bound DNA (shifted) fraction. The amount of DNA at plateau was determined by comparison to DNA standards using a PhosphoImager (Molecular Dynamics). Active protein concentrations were thus expressed in DNA equivalents (see Chadwick et al., 1970).


f

different experiments. Under these condition the K_d could easily be determined by the amount of active protein needed to occupy half the DNA binding sites (see Riggs et al., 1970 and Johnson et al., 1979). This follows from the fact that starting with the equations,

$$\text{protein-DNA} \leftrightarrow \text{protein}_{free} + \text{DNA}_{free},$$

$$K_d = [\text{protein}_{free}][\text{DNA}_{free}]/[\text{protein-DNA}]$$

and,

$$\text{DNA}_{total} = \text{DNA}_{free} + \text{DNA}_{bound}$$

it can be shown that,

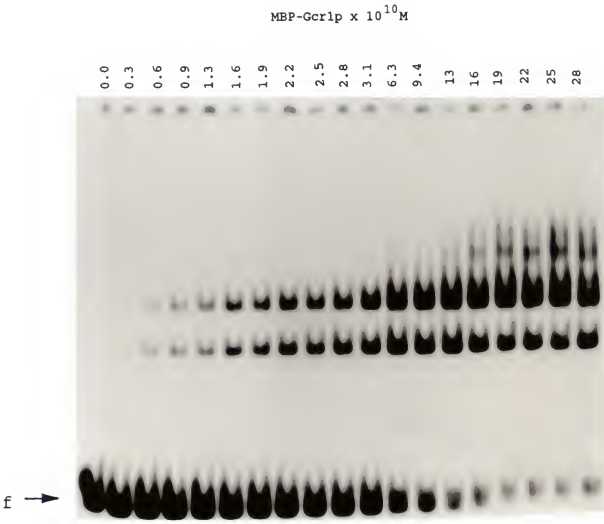
$$K_d = [\text{protein}]_{1/2} - 1/2[\text{DNA}_{total}]$$

where $[\text{protein}]_{1/2}$ is the concentration of protein when $[\text{DNA}_{free}]$ equals $[\text{DNA}_{bound}]$ (Riggs et al., 1970). Therefore when $[\text{DNA}_{total}]$ is much less than K_d the $[\text{protein}]_{1/2}$ gives a good estimate of the K_d.

A typical titration experiment is shown in Figure 11. A graphical representation of the date is displayed in Figure 12. Following the example of Letovsky and Dynan (1989), the total shifted complexes were considered bound complexes and were combined in determining the K_d. Quantitatively, the assay was essentially treated as a filter binding experiment (see Letovsky and Dynan, 1989). From these experiments the calculated apparent K_d of the Gcr1p DNA-binding domain with its recognition site was 2.9×10^{-10} M.

At higher protein concentration additional bands appeared with decreased electrophoretic mobility (Figure 11, lanes 15-19). At concentrations above 20 nM all the

Figure 11. Determination of Gcr1p DNA-binding domain binding affinity. A single Gcr1p binding site (probe TPI_1 in Table 3.) was radiolabelled and held at a concentration of 1.9×10^{-11}M. Increasing amounts of purified MBP-Gcr1p(6900-844) fusion-protein were incubated with the DNA probe. Samples were run into a 5 percent non-denaturing fel as described in Material and Methods. Numbers above the lanes represents the concentration of the DNA-binding activiy.


MBP-Gcr1p x 10^{10}M
0.0
0.3
0.6
0.9
1.3
1.6
1.9
2.2
2.5
2.8
3.1
6.3
9.4
13
16
19
22
25
28
f

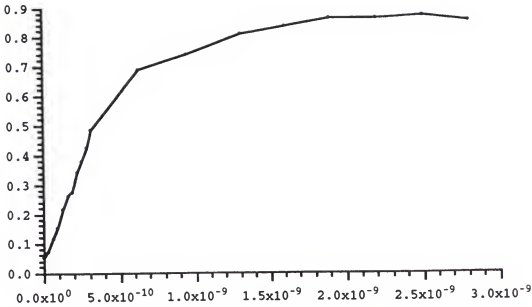


Figure 12. Graphical representation of binding affinity. The ratio of DNA_{bound}/DNA_{total} is plotted against increaseing concentration of MBP-Gcr1p(690-844). Quantitation of data from Figure 11. was achieved by the use of a phospoImager (Molecular Dynamics). The appearant K_d of the MBP-Gcr1p-DNA interaction corresponds to the concentration of active protein when DNA_{bound}/DNA_{total} is 0.5 (Riggs et al., 1970).

complexes were in the higher bands (data not shown).

Specificity of the Gcr1p DNA-Binding Domain

The results of the binding affinity studies showed that Gcr1p had a higher affinity for DNA than had been expected. Since nonspecific DNA-binding appeared to occur at less than 100-fold of the measured K_d, this suggested that the difficulty in demonstrating DNA-binding activity may have been due to a low specificity of binding. Specificity of binding is defined as the ratio of binding to a known binding site compared to binding to a random sequence of DNA (Affolter et al., 1990; Giese et al., 1991; Ferrari et al., 1992). Most assays for Gcr1p binding did not use purified Gcr1p (Baker, personal communication) and therefore poly dI/dC DNA was used as an competitor of nonspecific binding. It was possible, that if Gcr1p had a low specificity, this could mask the binding of Gcr1p. To test this I next set out to determine the specificity of binding of the Gcr1p DNA-binding domain to its recognition site.

To measure the specificity of Gcr1p for its high-affinity binding site in UAS_{TPI1} a series of competition experiments were performed. Competition experiments were performed with either specific DNA containing a known Gcr1p-binding site, or nonspecific DNA containing random DNA. The specific competitor, containing a Gcr1p-binding site, was made by annealing the complementary oligonucleotides HB39 and HB40 (see Table 1). The nonspecific competitor was made by extending oligonucleotide HB57 with HB58 (see Table 1). This DNA contained a random sequence at its core. A typical

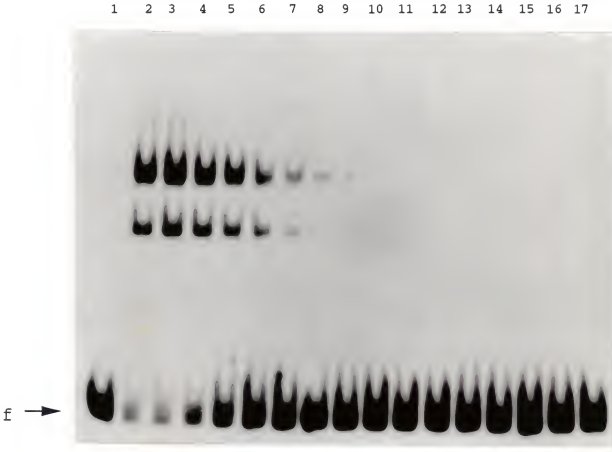


Figure 13. Competition experiment using specific competitor. DNA band-shift experiments were performed as described in Material and Methods. Radiolabeled DNA probe and purified MBP-Gcr1p(690-844) were held constant and increasing amounts of specific competitor were added to mixtures.

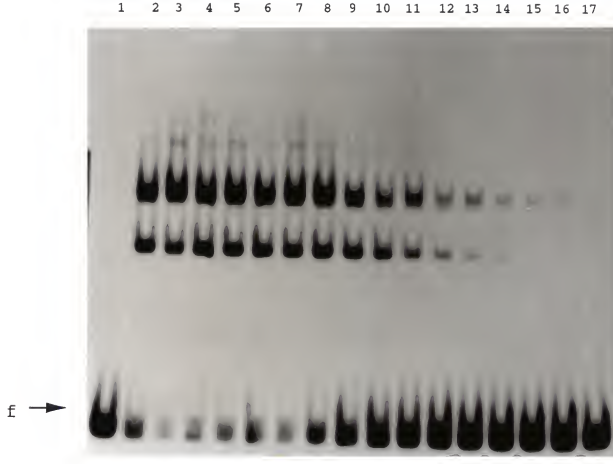


Figure 14. Competition experiment using non-specific competitor. DNA band-shift experiments were performed as described in Material and Methods. Radiolabeled DNA probe and purified MBP-Gcr1p(690-844) were held constant and increasing amounts of specific competitor were added to mixtures.

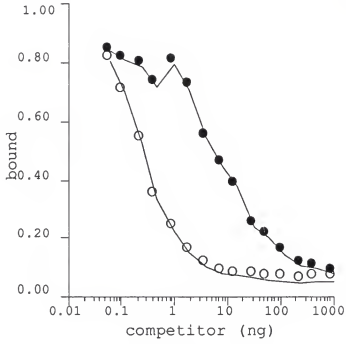


Figure 15. Graphical representation of competition experiments. Amount of competitor added (in nanograms) is depicted on horizontal axis. Specific competitor is represented by open circles; non-specific competitor is represented by closed circles. Vertical axis is ratio of DNA shifted with competitor against no competitor.

experiment is shown in Figure 13. An increasing amount of unlabeled competitor DNA containing a known Gcr1p-binding site was incubated with fixed concentrations of radiolabeled DNA and fusion-protein. In Figure 14 an increasing amount of unlabelled DNA containing random sequence was used as competitor. A graphical representation of these experiments is presented in Figure 15. From these experiments the difference in the amount of competitor necessary for competition of one-half binding is approximately 33-fold.

Bending of DNA in the Gcr1p-DNA Nucleoprotein Complex

Over the last few years a growing number of regulatory proteins which bind to DNA have been shown to induce bending of DNA. X-ray crystallographic data clearly reveals that the physical structure of DNA is contorted when bound to some of the best characterized transcriptional factors. For example, as determined by 3 angstrom resolution, the catabolite activator protein (CAP) induces a 90 degree bend in DNA by forming two 40 degree kinks around the dimeric protein (Schultz et al., 1991). An increasing body of literature has revealed that DNA bending plays an important role in transcriptional regulation, as well as control of replication (Williams et al., 1988; Doepsel and Khan, 1986; Zahn and Blattner, 1985; Mukherjee et al., 1985; Ryder et al., 1986; Snyder et al., 1986) and recombination (Better et al., 1982).

Protein-induced DNA bending can be detected easily by polyacrylamide gel electrophoresis (Calladine et al., 1991), and a technique to detect, map, and measure the extent of bending has been developed (reviewed in Crothers et al.,

1991). The technique, called comparative electrophoresis circular permutation assay, exploits the finding that when a protein is bound to DNA the shape of the DNA affects the mobility of the nucleoprotein complex to a large degree in nondenaturing gel electrophoresis (Wu and Crothers, 1984). The strategy is illustrated in Figure 16. By varying the position of the DNA-binding site along a DNA fragment of constant length, differing shapes of the nucleoprotein can form with corresponding degrees of electrophoretic mobilities. This is usually achieved by cloning the DNA-binding sequence into the middle of a tandem dimer containing multiple restriction sites in a specially design vector. After cleaving with the different enzymes, fragments of equal size but with the binding site in different positions are prepared as probes for DNA band-shift assays.

As mentioned earlier, known Gcr1p- and Rap1-binding sites are juxtaposed in the UAS of the glycolytic genes. Rap1p has been shown to bend its DNA recognition site in the UAS of the *rpg-1* gene (Vignais and Sentenac, 1989). *In vivo* Edward Scott has shown that the Rap1p-binding site is occupied in the absence of Gcr1p (Huie et al., 1992; Scott and Baker, 1993). Therefore, it was reasoned that the DNA-binding site recognized by Gcr1p *in vivo* may be in a bent state. If this is the case, then the true substrate for Gcr1p-binding may in fact be bent DNA. Kahn and Crothers (1992) have shown that the affinity of CAP for bent DNA is 200-fold greater than for linear DNA. If the target of Gcr1p

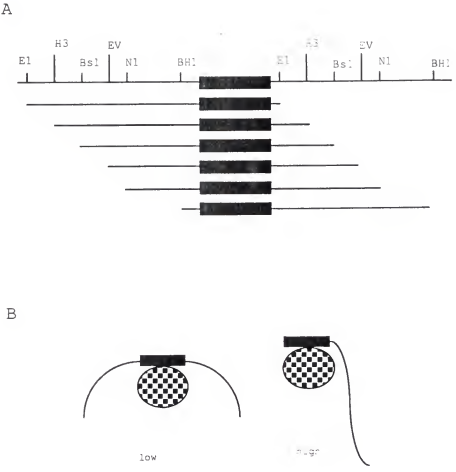


Figure 16. Circular permutation assay for determination of DNA bending A) A cloned tandem dimer is cleaved with a series of restriction enzymes generating probes of equal length but with varing DNA-binding site postitions. B) When a protein binds to a site in the middle of the fragment, producing a bend in the center, the DNA is contorted to a lower electrophoretic mobility form. When a protein binds to a site near the end of the fragment, the DNA is less cotorted producing a fragment with higher electrophoretic mobility (Figure based on Crothers et al, 1991 and Prentki et al., 1987).

is bent DNA, then it can be detected by the circular permutation assay.

I set out to test first if Rap1p also bends DNA when bound to its binding-site in the UAS of a glycolytic gene. Rap1p was translated in rabbit reticulocyte lysates and incubated with radiolabeled DNA from UAS_{TPI1} (a schematic of the probe is shown in Figure 17.) under standard band-shift assay conditions. Samples were then run on a 5% TE non-denaturing PAGE which had been pre-run for 1.5 hours. Results are shown in Figure 18. When the Rap1p-binding site is located in the middle of the DNA fragment (restriction endonucleases *EcoRV* and *NheI*) the mobility is decreased relative to the fragments containing the Rap1p-binding site at the ends of the fragments (restriction endonucleases *EcoRI* and *BamHI*). This result is consistent with Rap1p-inducing DNA bending.

The purified MBP-Gcr1p(690-844) fusion protein was used to assay if the Gcrp1p DNA-binding domain bends DNA in the nucleoprotein complex. Protein was incubated with radiolabeled DNA under standard DNA band-shift assay conditions. Samples were then run into a 10% nondenaturing 0.5 x TBE PAG which had been pre-run for 1.5 hours. The results are shown in Figure 19. When the Gcr1p-binding site is located in the middle of the fragment (restriction endonucleases *EcoRV* and *NheI*) the mobility is decreased relative to the fragments containing the Gcr1p-binding site at the ends of the fragments (restriction endonucleases *EcoRI* and *Bam*HI). These results are consistent with the notion

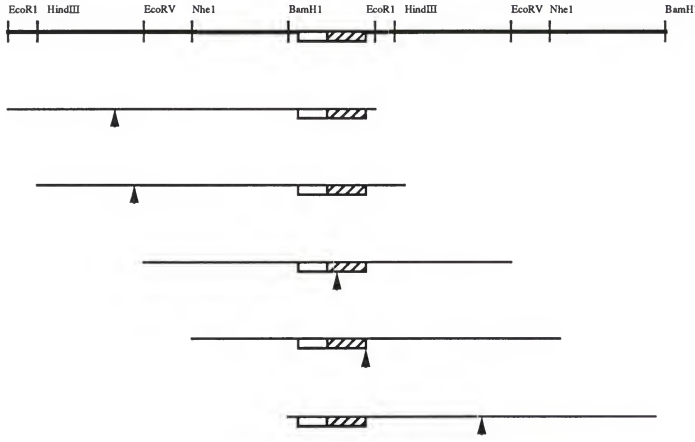


Figure 17. Schematic representation of the probes used in the circular permutation assay. The box represents the DNA fragment cloned into vector pCY4. Hatched segement represents sequence corresponding to the UAS of *TPI1* (see Table 3.). Digestion with the indicated enzymes produce fragment of the same molecular weight with the UAS of *TPI1* in varying positions. Arrow head signifies the center of the probe.

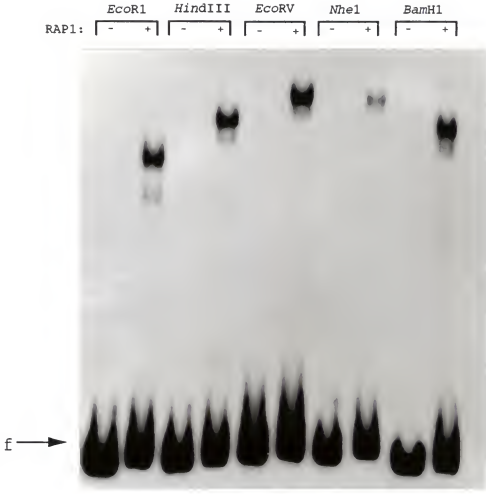


Figure 18. DNA is bent in the Rap1p-DNA nucleoprotein complex. Circular permutation assay was performed as described in Material and Methods. Schematic of the DNA probes used in this assay is shown in Figure 17. Two microliters of Rap1p polypeptide translated in rabbit reliculocyte lysates were incubated with radiolabed probes and then run into a 5 per cent non-denaturing polyacrlamide gel. (+) represents addition of Rap1p; (-), absence of Rap1p. f, free probe. Enzymes listed above lanes indicates probes shown in Figure 17.

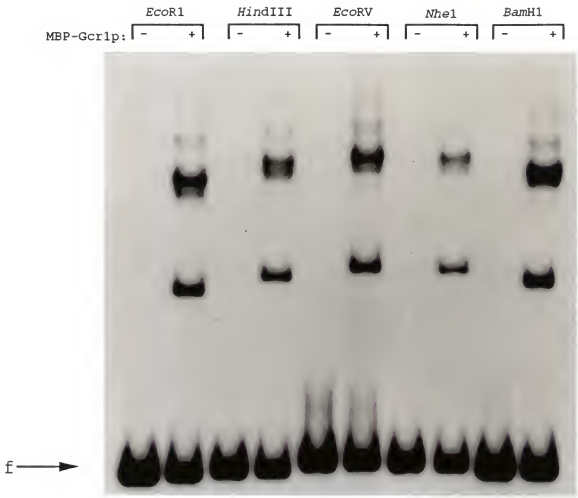


Figure 19. DNA is bent in the Gcr1p-DNA nucleoprotein complex. Circular permutation assay was performed as described in Material and Methods. Schematic of the DNA probes used in this assay is shown in Figure 17. One microliters of purified MBP-Gcr1p(690-844) fusion-protein were incubated with radiolabed probes and then run into a 5 per cent non-denaturing polyacrlamide gel. (+) represents addition of MBP-Gcr1p(690-844); (-), absence of MBP-Gcr1p (690-844). f, free probe. Enzymes listed above lanes indicates probes shown in Figure 17.

that Gcr1p does indeed contort the DNA when bound, or that bent DNA is a more favorable target for Gcr1p (see Discussion).

DISCUSSION

> For skepticism is this: that an unknown quantity, some *x*, can explain everything. But when everything is explained through an *x* which is not explained, then *in toto* nothing is explained, nothing at all. If this is not skepticism, then it is superstition.
>
> Søren Kierkegaard (p.251)

> Since Copernicus man has been rolling from the center toward *x*.
>
> Friedrick Nietzsche (p.8)

The enzymes of the glycolytic pathway constitute approximately 30-60 percent of the soluble cellular proteins in *Saccharomyces cerevisiae* (Hess et al., 1969 and Fraenkel, 1982). The observation that the most abundant mRNA species in this organism code for glycolytic enzymes (Holland and Holland, 1978), and the demonstration of high-level expression of heterologous genes using glycolytic gene promoters (Bitter and Egan, 1984; 1988; Bitter et al., 1987) revealed that these promoters are among the most powerful known in any species. The promoters of the glycolytic genes have been exploited commercially to manufacture recombinant human erythropoietin and recombinant hepatitis B viral antigens (Bitter and Egan, 1984; 1988; Bitter et al., 1987).

The upstream activating sequence, UAS, elements of many glycolytic enzyme genes have been studied in detail and a

common theme has emerged. A growing number of factors, including Reb1p, Rap1p, Abf1p, Gcr1p, Gcr2p, and Gal11p, appear to be the first proteins to assemble at these UAS elements, effect chromatin structure, and activate transcription (see Scott and Baker, 1993).

Of the DNA-binding proteins known to bind in the UAS elements of glycolytic genes, the binding of Rap1p and Gcr1p appears to be crucial for UAS activity. A role of Gcr1p in glycolytic gene expression was first shown when mutations were isolated in *GCR1* which resulted in a 20-fold reduction in the expression of most of the glycolytic enzymes genes (Clifton et al., 1978). Baker (1991) showed that full-length Gcr1p translated *in vitro* binds to the CTTCC sequence element in the UAS of *TPI1*. Mutations in CTTCC sequence elements in the UASs of the genes *TPI1* (Scott and Baker, 1993), *TDH3* (Bitter et al., 1991), *PGK* (Chambers et al., 1988), *ENO1*, and *PYK* (Buchman et al., 1988) impaired the ability of the UAS elements to drive expression of reporter genes.

RAP1 is an essential gene (Shore and Nasmyth, 1987) with pleiotropic functions. Rap1p is known to act as both an activator and repressor of transcription (Shore and Nasmyth, 1987), and it plays a role in the maintenance of telomeric structure (Buchman et al., 1988; Longtine et al., 1989; Lustig et al., 1990; Conrad et al., 1990). A role for Rap1p in the activation of glycolytic gene expression was first suggested when Rap1p DNA-binding sites were noted in the 5' noncoding region of many glycolytic genes (Capieux et al., 1989). Mutation-analyses of the Rap1p DNA-binding sites in

the UAS elements of *TPI1* (Scott et al., 1990), *TDH3* (Bitter et al., 1991), *PGK1* (Chambers et al.,1989), *ENO1* (Machida et al., 1989; Brindle et al., 1990), *ENO2* (Brindle et al., 1990), *PYK1* (McNeil et al., 1990), *PDC1* (Butler et al., 1990), and *ADH1* (Tornow and Santangelo, 1990) has shown that Rap1p-binding is essential for full activity of these UASs. In each case mutation of the Rap1p-binding site reduced expression more than ten-fold. It has additionally been shown that the Rap1p-binding site from *PGK* alone is unable to confer UAS activity to a test promoter (Stanway et al., 1989). Thus, whereas both Gcr1p and Rap1p DNA-binding sites are required for high-level glycolytic gene expression, neither site alone is sufficient for high-level expression.

The experiments in this study demonstrate conclusively that Gcr1p is a DNA-binding protein, and that Gcr1p interacts with the CTTCC sequence element directly. Additionally a Gcr1p consensus-sequence is approximated, and an initial characterization of the DNA-protein interaction is presented. Although this study focused mainly on Gcr1p, the results bear strongly on how Gcr1p and Rap1p interact, and how they exert a synergistic effect at the UAS of the glycolytic genes. One feature appears to be common to all of the glycolytic gene promoters regulated by *GCR1*: the finding of adjacent Gcr1p- and Rap1p-DNA-binding sites (Huie et al., 1992). The implications of the adjacent binding sites will be discussed in the context of the findings of this study.

The DNA-binding Domain of Gcr1p

Two lines of evidence identify a DNA-binding domain in the carboxy-terminus of Gcr1p. First, carboxy-terminus truncations of Gcr1p abolish DNA binding. Second, all fusion-proteins containing the carboxy-terminal 154 amino acid residues of Gcr1p retained sequence-specific DNA-binding activity. Additionally, the MBP-Gcr1p(690-844) fusion-protein--containing the carboxy-terminal 154 amino acid residues--protected from *in vitro* DNaseI digestion CT-boxes which are protected in a *GCR1*-dependent manner *in vivo* (Huie et al., 1992).

It had been previously suggested, on the basis of DNA sequence analysis, that the carboxy-terminus of Gcr1p may contain a DNA-binding domain since a possible helix-turn-helix (H-T-H) motif was identified at amino acids 784 to 803 (Baker, 1986). The H-T-H structure forms the basis of a DNA-binding domain in a superfamily of DNA-binding proteins (Harrison and Aggarwal, 1990). The Gcr1p fusion-proteins that contained amino acids 706 to 844 and amino acids 784 to 844 failed to bind DNA even though they contained the putative H-T-H. This negative result does not exclude the H-T-H motif from forming the core of the actual DNA-binding domain, as smaller deletions could have disrupted the domain.

In an effort to compare the Gcr1p DNA-binding domain to other proteins, sequence analysis was performed on the entire protein with particular attention paid to the carboxy terminus. Comparison against updated databases using the algorithms BLAZE (Smith and Waterman, 1992), BLAST (Altschul

et al, 1990), and FASTA (Pearson and Lipman, 1988) [with a K_{tup} of 1 for increased sensitivity (Pearson, 1991)], uncovered no obvious homology to other known DNA-binding proteins.

A number of sequence-specific DNA-binding proteins which recognizes a centrally located 5'-GGAA-3' core (which is complementary to 5'-TTCC-3') contain the *ets*-domain (Karim et al., 1990). The primary amino acid sequence of the Gcrlp DNA-binding domain displays no obvious homology to this class of proteins. Although there are a tyrosine and two tryptophan residues spaced 15 and 18 amino acids apart similar to PU.1, a member of the *ets* family of proteins, Gcrlp lacks the highly-conserved amino-acid residues characteristic of the *ets*-family of proteins.

Although sequence analysis failed to identify other proteins that contain a potential Gcrlp DNA-binding domain, the algorithm of Karlin and Brendel (1992)--Statistical Analysis of Protein Sequences (SAPS)--performed on the entire Gcrlp amino acid sequence correctly predicted a DNA-binding domain in the region of Gcrlp to where it was functionally mapped (see Figure 20). This algorithm scores amino acids based on their frequency in known DNA-binding domains. So, although the Gcrlp DNA-binding domain does not show obvious homology to known DNA-binding domains, the amino acids residues present in the domain are well represented in known DNA-binding domains.

Although no obvious homologies were found with known DNA-binding proteins, when the DNA sequence coding for this

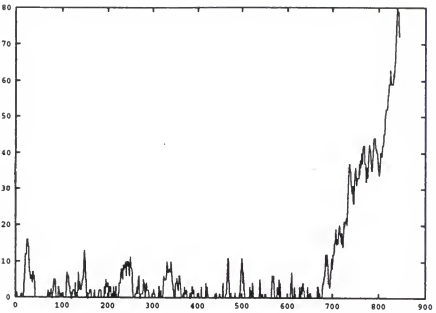


Figure 20. SAPS analysis of Gcr1p. The algorithm of Karlin and Brendeal (1992) performed on the entire Gcr1p amino acid sequence predicted a DNA-binding domain in the carboxy-terminal region of Gcr1p. Numbers on the vertical axis represent E_i, the excursion score. Numbers on the horizontal axis represents the amino acid residues of Gcr1p. A significance threshold of 5% corresponds to an E_i score of 26. The carboxy-terminal region of Gcr1p has a significantly high E_i score. This region corresponds to the of DNA-binding activity mapped in this study. The amino acid matrix presented in Karlin and Brendel (1992), with a yeast codon bias, was used to calculate E_i for this analysis.

region was used as a probe to screen a genomic DNA from a variety of organisms from yeast to humans, homologous sequence could be detected by low stringency hybridization (Huie and Baker, unpublished observations). Thus, homologous Gcr1p DNA-binding domains may not yet have been cloned and reported.

A DNA Consensus Sequence for Gcr1p Binding

In this study it was shown that Gcr1p can interact directly with the CTTCC sequence elements found in the upstream regions of *TPI1*, *PGK*, *ENO1*, *PYK*, and *ADH1*. From these experiments a degenerate consensus Gcr1p-binding site was formulated:

(T/A)N(T/C)N(G/A)N **C (T/A) T C C** (T/A)N(T/A)(T/A)(T/G)

The extent of the proposed consensus Gcr1p-binding site (16 nucleotides) was chosen on the basis of the area of protection observed in the *in vitro* DNase I footprinting study. This sequence is in good agreement with the GPE (for GRF1 [RAP1] site Potentiator Element) motif recently defined by Bitter et al. (1991) as:

G(A/**C) (A/T) T C C** (A/T)

The GPE motif, found adjacent to Rap1p-binding sites in glycolytic genes has no UAS activity by itself, but potentiated the activation effects of a Rap1p-binding site when placed next to it either upstream or downstream and in either orientation (Bitter et al., 1991). Bitter et al. (1991) concluded that Rap1p-binding sites and GPE are responsible for the majority of activation potential of the UAS of *TDH3* and suggested that GPE represented a binding site

for a protein. From the studies presented here, it is almost certain that GPE sites are Gcr1p-binding sites.

From the consensus Gcr1p-binding site listed above it was predicted (Huie et al., 1992) that Gcr1p would bind at position -451, relative to the start codon, in the UAS of *ENO2*. This site has recently been shown to bind Gcr1p *in vitro* (Willett et al., 1993).

It is appreciated that Gcr1p may interact with sequences which contain several mismatches to the proposed consensus sequence. In fact, it is likely that such cases will be found, since this consensus is base on a limited set of binding sites. Additionally, the true binding site *in vivo* may be affected by other proteins bound at adjacent sites (see below) or by the topology of the DNA in and around the Gcr1p-binding site. For example, the CTTCC sequence is often centered between a cluster of T and A residues spaced by 10 base-pairs (see, in particular, the sequence of $TPI1_1$, $TPI1_2$, and *ADH1* listed in Table 3). Runs of three or four A/Ts regularly spaced at 10 bp (one helical turn) are associated with intrinsically bent DNA (Wu and Crothers, 1984). Sequences involved in nuclear matrix attachment have patterns characteristic of intrinsically bent DNA (Anderson, 1986). This is worth noting since Rap1p purifies with the nuclear matrix (Cardenas et al., 1990) and since all known Gcr1p-binding sites are adjacent to Rap1p-binding sites (see below). Similarly, the B element of the ARS sequence, to which Abf1p binds, has been shown to have a strong intrinsic

bend (Williams et al., 1988; Eckdahl and Anderson, 1987; Snyder et al., 1986).

These points are reviewed to underscore the fact that proteins can recognize DNA without directly interacting with the bases. This type of recognition has been referred to as "indirect readout" (Otwinowski et al., 1988) or analogue recognition (Drew and Travers, 1985; Travers, 1989). For example, the *trp* repressor, as determined by X-ray crystallography, binds to its site by contacting the phosphate-backbone, without contacting any base-pair directly (Otwinowski et al., 1988). It has been observed that the H-T-H motif identified in the primary sequence of *GCR1* most closely resembles the H-T-H of the *trp* repressor (Baker, personal communication). If Gcr1p binds in such a manner, the primary DNA sequence of the consensus site may not accurately reflect the true determinants of binding. However, G residues are protected from methylation by DMS (Huie et al., 1991) which suggests a very close association between Gcr1p and the N^7 position of those G residues.

The Binding Affinity of the Gcr1p DNA-Binding Domain

Using a band shift assay with purified fusion-protein containing the DNA-binding domain of Gcr1p it was here shown *in vitro* that the Gcr1p DNA-binding domain binds with a relatively high affinity (K_d = 2.9 x 10^{-10} M) to one of the sites it recognizes *in vivo*. Objections may be raised to the fact that these studies were performed with fusion-protein and that only the DNA-binding domain and not the full-length Gcr1p polypeptide was used. Unfortunately, the instability

of full-length Gcr1p makes the full-length polypeptide difficult to use for these types of measurements. As mentioned in the introduction, transcriptional activators often contain "disorganized" structures, such as exposed hydrophobic residues, which undergo an induced-fit with other factors when they interact (reviewed in Frankel and Kim, 1991). This disorganized structure makes transcriptional activators very sensitive to proteases. Full-length Gcr1p seems to be unstable whether translated in rabbit reticulocyte lysate or manufactured in *E. coli*.

However, due to the modular nature of transcriptional activators, many structural questions can still be addressed and have been addressed by studying the discrete domains. For example, the RNA-binding affinity of a short peptide that contains the RNA-binding domain of the HIV tat protein gives the same measurements as with the intact protein (Calnan et al. 1991; Dingwall et al., 1990). The DNA binding-domain of LEF-1 expressed as a peptide or a GST-LEF-1 fusion protein, containing only the DNA-binding domain, gives similar binding affinities (Giese, 1991). Nevertheless, the binding affinity from the studies presented here must be understood in the context of studying the DNA-binding domain, and not the full-length protein.

The binding affinity of the DNA-binding domain of Gcr1p falls into the middle of a range of affinities described for known DNA-binding domains of transcriptional activators. A few factors are listed in Table 4. for comparison.

Table 4. DNA-Binding Affinity of Select Transcriptional Activators

factor	K_d	reference
CAP	3×10^{-8}	(Liu-Johnson et al., 1986)
Antennapedia	$1 - 2 \times 10^{-9}$	(Affolter et al., 1990; Corsetti et al., 1992)
Sp1	5.3×10^{-10}	(Letovsky and Dynan, 1989)
Rap1p	1.3×10^{-11}	(Vignais et al., 1990)
NF-1	2.1×10^{-11}	(Eberly et al., 1985)
ftz	2.5×10^{-11}	(Florence et al., 1991)
HSTF	4×10^{-12}	(de Vries and Koogh-Schuuring, 1973)

Vignais et al. (1990) measured the apparent K_d of Rap1p for its optimal binding site at 1.3×10^{-11} M. Comparing this value with that of Gcr1p gives good agreement with band shift data from other studies (Baker, 1991; Huie and Baker, unpublished data), where it was estimated that 50-fold more full-length Gcr1p had to be added, compared to Rap1p, to see the same degree of DNA-binding of a labeled fragment containing both binding sites.

In the DNase I footprint assay presented in this study (Figure 8), only two specific regions in a 234 bp fragment were protected from DNase I digestion. The regions are identical to genetically defined *cis*-elements responsible for high-level expression of *TPI1* (Scott and Baker, 1993). Furthermore, Edward Scott demonstrated *in vivo* that these sequences are protected in wild-type strains and deprotected in *gcr1* mutant strains (Huie et al., 1992; Scott, 1992; Scott and Baker, 1993).

One unanticipated result of this study was the relatively high-affinity but modestly low-specificity of the interaction of the Gcr1p DNA-binding domain with DNA when measured *in vitro*. Other lines of evidence also support the notion that Gcr1p has a low specificity for its specific binding site. Preliminary results using a PCR-based oligonucleotide selection and amplification technique suggests that the Gcr1p DNA-binding domain is very poor at enriching for highly specific sites (Esiner and Baker, unpublished data). Theoretical considerations (Irvine et al., 1991) of the technique are consistent with the

observations found from experiments (Huie and Eisner, unpublished observation) that Gcr1p has a low specificity. The low specificity was unanticipated because even though Gcr1p is expressed in low amounts in the cell, its sites at UAS_{TPI} seem to be fully occupied *in vivo* (Scott and Baker, 1993; Scott and Baker, unpublished data).

In this respect, the Gcr1p DNA-binding domain can be added to a growing list of transcriptional factors that posses a high affinity for DNA, yet a low specificity of interaction with their specific binding sites. Other characterized transcription factors with these properties are found in humans, mice, insects, and bacteria. For example, the homeodomain of *Antennapedia* from *Drosophila* has a high affinity for DNA, but displays a specificity of less than 100-fold (Affolter et al., 1990). As measured by kinetic studies, the glucocorticoid receptor complex with its DNA binding site displays only about a 10-fold difference in affinity for specific verse nonspecific DNA (Schauer et al., 1989). The transcriptional activator LEF-1, which binds to the T-cell receptor enhancer, has a measured dissociation constant of 10^{-9} M, yet a specificity of only 20-40 fold (Giese, 1991). The mammalian sex-determining gene, SRY, also encodes for a DNA-binding protein with a very low specificity (Nasrin et al., 1991; Harley et al., 1992). All of these proteins, despite the low specificity, demonstrate clear areas of protection in DNase I footprinting studies. In addition, many of these proteins, like Gcr1p, have been shown

to bend DNA upon binding (Giese et al., 1992; Ferrari et al., 1992).

The low specificity, seen in many transcriptional *activators*, can be contrasted with the properties observed in some transcriptional *repressors*. The λ repressor binds to its specific operator sequences with affinities as high as 10^{-13} M and binds to specific DNA with a 500,000-fold higher affinity than to nonspecific DNA (Sauer et al., 1990; Frankel and Kim, 1991). Since it is specificity that is more important for assembling transcriptional complexes, a number of implications can be drawn from questions that arise from the finding of transcriptional activators which have a relatively high affinity for DNA and yet a low specificity.

An immediate question arises as to how proteins with a low specificity find their DNA-binding sites in their genomes. This is a serious question which cannot be dismissed with a flippant reference to skepticism. Possible models are testable and may provide clues as to how other proteins acting at UASs, which do not posses activation domains, exert their effects. It is worth reviewing at this point possible mechanism by which proteins with low specificity can find their binding sites *in vivo*.

One possible mechanism is simply by high-level expression of the protein. This has been proposed to be the manner in which integration host factor (IHF) finds its site (Giese, 1991). Although IHF is a sequence-specific DNA-binding protein (Craig and Nash, 1984), it also binds to nonspecific DNA (Bonnefoy and Rouviere-Yaniv, 1991). IHF is

estimated (Giese et al., 1991) to be in high abundance in the cell (20,000 molecules per cell), and this may drive the factor to its site by mass action. IHF plays many different roles in the cell (reviewed in Friedman, 1988), but it is believed, fundamentally, to be a structural or 'architectural' protein. This mechanism of high-level expression is probably not the mechanism used by transcriptional activators. A transcriptional activator with a low specificity may not be tolerated well by the cell when expressed at a high level. In fact many, if not most, transcriptional factors are toxic to their host when expressed at high levels. Presumably this is due to either nonspecific binding or squelching of auxiliary transcription factors (Gill and Ptashne, 1988). Full-length Gcr1p in fact seems to be toxic when expressed at high levels in yeast (Holt and Baker, unpublished observation).

Another way in which proteins with low specificity can find their binding sites is by interacting with other proteins. According to this view, the K_d for a specific site would be decreased by a cooperative interaction with another DNA-binding protein (Giese et al., 1991). The *in vivo* binding site would then consist of a nucleoprotein complex. The other DNA-binding protein(s) would presumably have a stronger DNA-binding affinity and specificity and bind to its site(s) first. This complex would then facilitate the binding of the protein with a low specificity through a protein-protein interaction. This mechanism has been proposed for the facilitation of the factor OTF-1 by Sp1

(Janson and Pettersson, 1990). Using band-shift assays these authors demonstrated a 10-fold increase in affinity of OTF-1 to its site in the presence of Sp1 when an Sp1 site was adjacent to an OTF-1 site. Furthermore, OTF-1 binding was decreased when a 15 bp segment was inserted between the Sp1 and OTF-1 binding sites (Janson and Pettersson, 1990). In perhaps a more classic example, the λ repressor has been shown to facilitate the binding of another repressor to a lower affinity site by a protein-protein interaction (Hochschild and Ptashne, 1986).

An alternative mechanism, which does not exclude the one mentioned above, is one in which the specificity of a protein-DNA interaction is increased by another protein modifying the DNA topology. Wu and Crothers (1984) suggested that protein induced bending of DNA can facilitate additional protein-DNA interactions. In this scenario, the DNA sequence is contorted making a more favorable binding site and thus increasing the affinity. This contortion could be caused by the binding of a protein to an adjacent site which then distorts the DNA making a more favorable binding site for the protein with low specificity; or alternatively, a primary DNA sequence could alter the adjacent binding site to produce a more favorable site. An example of the former mechanism was proposed (Flashner and Gralla, 1988) for the stimulation of CAP binding to *lac* DNA by HU - the major histone like protein of *E coli*. For an example of the latter mechanism, it has been shown that the affinity of CAP for its binding site is increased 200-fold when it is located on an intrinsically

bent piece of DNA (Kahn and Crothers, 1992). Likewise, IHF has also been shown to have a greater affinity for curved DNA (Bonnefoy and Rouviere-Yaniv, 1991). This mechanism, in which a protein contorts a nearby site, was proposed for the functional versatility of Rap1p when it was first noted that Rap1p bends DNA (Vignais and Sentenac, 1989).

A third mechanism in which specificity could be increased would be due to modification of the DNA-binding protein. For example, the binding of sequence-specific transcription factors to DNA can be affected either positively or negatively by phosphorylation (reviewed in Hunter and Karin, 1992). Examples of negative regulation include the products of *c-myb* and *c-jun* in which phosphorylation decreases the protein's DNA-binding affinity (Luscher et al., 1990; Boyle et al., 1991). Myb binds to several characterized Myb-response-elements (MRE) with varying affinities. Phosphorylation of Myb completely inhibits binding to the MREs with low affinities, but is not very efficient at inhibiting Myb binding to MREs with high binding affinity, thus in effect increasing the specificity (Luscher et al., 1990). The factor SRF, on the other hand, increases its affinity for DNA when phosphorylated at serine residues (Janknecht et al., 1992; Maris et al., 1992). This is presumably due to a conformational change of the protein when phosphorylated (Marak and Pryuses, 1991).

In the final analysis, transcription complexes are held together by multiple cooperative interactions, even if some individual interactions are of low specificity. As Frankel

and Kim (1991) have pointed out, it would be difficult to dissociate or regulate the activity of a transcription complex if every component had an exceedingly high affinity or specificity for every other component; and furthermore, extremely tight or specific interactions might interfere with the combinatorial use of factors by many promoters (see Dynan, 1989). Ribosomes have been suggested as an apt analogy to the assembly of transcription complexes (Frankel and Kim, 1991). In general, both systems are complex macromolecular assembles, millions of kiloDaltons in size, built upon multiple cooperative protein-protein and protein-nucleic acid interactions of varying affinities and specificities. Only when the entire complex is assembled is the true *specificity* of the system revealed. Individual interactions with modest specificity seems to be inherent to designs of systems with such complexity--with cooperativity resulting largely from entropic factors (Creighton, 1983).

The Bending of DNA by Raplp and Gcrlp

It has been pointed out (Kerppola and Curran, 1991) that the circular permutation assay cannot distinguish between static DNA bends or other alterations, such as triple helix formation. This caveat is included with the term DNA "bend" when referring to the patterns of electrophoretic mobility observed in this study. With this in mind, it was shown here that Gcrlp bends DNA upon binding to it. It is important to note here that this distortion of DNA represents a thermodynamic barrier which if reduced would presumably facilitate binding (see Kahn and Crothers, 1992). This can

be represented as follows:

1. DNA_t <-> DNA_{bent} k_1

2. DNA_{bent} + Gcr1p <-> Gcr1p-DNA k_2

3. $k_t = k_1 k_2$

If DNA_{bent} is the true binding substrate (i.e. has a higher affinity) for Gcr1p then the equilibrium constant of equation 1. will determine the effective concentration of the DNA-ligand for Gcr1p. Factors that enhance the formation of DNA_{bent} would then enhance the formation of the Gcr1p-DNA complex. This concept of DNA bending facilitating binding of protein has been compared (Kahn and Crothers, 1992) to the improved binding of unwinding ligands to supercoiled DNA (Davidson, 1972).

A protein which may enhance reaction 1. presented above is Rap1p. Rap1p has been shown to bend DNA upon binding to the rpg box (Rap1p-binding site) from the UAS of *rpg-1* (Vignais and Sentenac, 1989). Here I have shown that Rap1p also bends DNA when bound to its recognition site in the UAS of *TPI1*. If Gcr1p binds with greater affinity to bent DNA, then Rap1p ability to bend DNA in and around the Gcr1p binding site may enhance the binding of Gcr1p. This mechanism for Rap1p was proposed by the authors who originally showed that Rap1p bends DNA to account for its functional versatility (Vignais and Sentenac, 1989). These authors also demonstrated that proteolytic fragments of Rap1p, which retained only the DNA-binding domain, were unable to bend DNA. It is possible that deletions of Rap1p which cannot bend DNA could effect Gcr1p binding if bending

of DNA plays a significant role in facilitating the binding of Gcr1p.

The Significance of Adjacent Rap1p and Gcr1p Binding Sites in the UAS of the Glycolytic Genes

Given the findings that Gcr1p has a low specificity of interaction with its specific DNA sites as measured *in vitro*, that Gcr1p either recognizes bent DNA or bends DNA upon binding, and the models discussed above, what can we surmise about the role of the other factors which have DNA binding sites in the UASs of the glycolytic genes and which are also required for high-level expression of the corresponding genes?

Due to the low specificity of Gcr1p interaction with its binding sites, the concept of facilitation most likely can explain the findings. The factor that most likely would play a role in facilitating the binding of Gcr1p to its sites is Rap1p. A number of observations that support this view will be reviewed at this point.

The first observation is the relationship between Rap1p- and Gcr1p-binding sites in the UAS of the glycolytic genes which are dependent upon *GCR1* for full expression. All known Gcr1p-binding sites (Willett et al., 1993; Huie et al., 1992; and this study) are found next to Rap1p-binding sites. Furthermore, putative Gcr1p-binding sites in the UAS of the *GCR1* dependent glycolytic genes reveal adjacent Rap1p-binding sites. The presence of a Rap1p-binding site near putative Gcr1p-binding sites in the glycolytic genes raises the possibility that either a direct interaction between these two proteins occurs or that Rap1p affects the structure of

the DNA is such a way to increase Gcr1p-binding at its site. However, in the Ty2 retrotransposon UAS and enhancer, several Gcr1p-sites have been mapped which do not have adjacent Rap1p-sites (Turek, 1994). DNA-binding protein sites are adjacent to the Gcr1p-sites, but the identiy of the proteins are currently unknown.

A second point to consider is that for the genes *TDH3* (Bitter et al., 1991), *PGK1* (Chambers et al., 1990), and *ADH1* (Tornow et al., 1993) a Gcr1p-binding sites by itself is unable to convey UAS activity. However, Baker and Scott (1993) have shown that if a lexA binding domain is tethered to full-length Gcr1p, activation of a *lex* operator::*GAL1*:*lacZ* reporter can be observed. This activation occurs in the absence of a Rap1p- binding site (Scott and Baker, 1993). This demonstrates that Gcr1p can activate transcription in the absence of Rap1p if bound to DNA. This result was recently confirmed by other investigators (Tornow et al., 1993). That Gcr1p can activate transcription when bound to DNA through a lex operator is consistent with the notion that the role of Rap1p is to facilitate binding of Gcr1p and that Gcr1p contains an activation domain.

A third point: Rap1p is an abundant sequence specific DNA-binding protein in yeast which is known to bind to its sites with a very high degree of specificity. Vignais et al. (1990) measured the apparent K_d of Rap1p for its optimal binding site and for nonspecific DNA. These values were found to be 1.3×10^{-11} M and 8.7×10^{-6} M respectively. Thus, the specificity of Rap1p binding to its site verse random DNA

is six orders of magnitude. In addition to this high degree of specificity, Rap1p is expressed in high abundance in the cell. Its level has been estimated to be 4,000 molecules/genome (Buchman et al., 1988). Since Rap1p has an extremely high specificity and is expressed at high levels, it has been argued that all Rap1p-binding sites should be occupied in actively growing cells (Vignais et al., 1990). Thus, Rap1p should bind its site in the absence of Gcr1p.

From *in vivo* footprint analysis it is known that the Rap1p-binding site in the UAS_{TPI1} is protected in the absence of the *GCR1* gene product (Huie et al., 1992; Scott, 1992; Scott and Baker, 1993). A safe conclusion from this data is that the *GCR1* product is not necessary for Rap1p binding at its site in UAS_{TPI1} *in vivo*. On the other hand, using temperature sensitive *rap1* mutants, Jeff Smerage (Baker, personal communication) has preliminary evidence from *in vivo* footprinting experiments that Rap1p binding is required for Gcr1p binding at UAS_{TPI1}. This implies that Gcr1p binds to its DNA-binding site only in the presence of Rap1p bound at an adjacent site.

All these observations are consistent with the view that the role of Rap1p is to facilitate the binding of Gcr1p to its binding site, and that Gcr1p then provides an activation domain. What might be the possible mechanism whereby Rap1p carries out this facilitation? The observation of "skewing" of protection of the RAP1 site seen in *in vitro* DNase I footprinting experiments implies that a structural change near the Rap1p-binding site is occurring in the minor groove

since DNase I recognizes a vary narrow range (Drew and Travers, 1985; Suck and Oefner, 1986; Lahm and Suck, 1991). This is consistent with the idea that Raplp-induced bending of DNA affects nearby DNA binding sites for other proteins.

Evidence also exist, however, for a direct Raplp-Gcrlp interaction. Tornow et al. (1993) showed that the *GCR1* gene product could be co-immunoprecipitated with Raplp when the two proteins were epitope tagged. These results suggest a protein-protein interaction occurs between Raplp and Gcrlp.

The role of the DNA-binding domain in the function of Gcrlp was recently called into question. Tornow et al. (1993) interpreted experiments from their laboratory to suggest that the DNA-binding domain of Gcrlp may be dispensable for transcriptional activation activity. This was intriguing because previous experiments had suggested that the DNA-binding domain of the *fushi tarazu* polypeptide (a homeodomain) could be deleted and that the truncated protein could still alter gene expression, presumably without binding directly to DNA (Fitzpatrick et al., 1992).

Tornow et al. complemented a *gcr1* mutant with plasmids containing 3'-deletions of the *GCR1* structural gene and noted the appearance of large colonies after ten days of growth on minimal medium containing glucose. However, these authors did not show that the ability to complement the *gcr1* growth defect was linked to the plasmids encoding the truncated forms of Gcrlp. This control is essential because second-site pseudorevertants of the *gcr1* mutant can appear as large colonies. Drazinic et al. (1994), in an attempt to verify

this observation, have shown that a plasmid containing a 3'-deletion of *GCR1* introduced into the *gcr1* mutant strain HBY4 (which contains a deletion of the entire *GCR1* structural gene) were incapable of complementing the slow-growth phenotype. These experiments, which were done under conditions to select for retained plasmids, failed to confirm the results of Tornow et al. (1994) Only the full length *GCR1* construct was capable of complementing HBY4. These result suggest that Tornow et al. (1994) were selecting second-site pseudorevertants of the *gcr1* mutant.

Tornow et al. (1994) also suggested that Gcr1p exerts its affects through Rap1p-binding site. However, a number of different laboratories have shown that mutations of CTTCC elements (Gcr1p-binding sites) results in severely reduced expresssion of the cooresponding gene (Bitter et al., 1991; Buchman et al., 1988; Chambers et al., 1988; Scott and Baker, 1993; Willet et al., 1993). Furthermore, Drazinic et al. (1994) have shown by *in vivo* footprinting that mutations in both Gcr1p-binding sites within UAS_{TPI} which reduce expression of the cognate gene by 67-fold do not affect the occupancy state of the adjacent Rap1p-binding site. If Gcr1p was recruited to the UAS element soley by protein-protein interactions with Rap1p, then one would not expect mutations in the Gcr1p-binding sites to affect expression of the cognate gene. The Data taken as a whole supports an interdependency between Rap1p and Gcr1p binding site at glycolytic enzyme gene UAS elements in mediating *GCR1*-dependent gene expression.

Conclusions

The experiments described in this study demonstrate that Gcr1p possesses a modular DNA-binding domain, located in the carboxy-terminal 154 amino acids, that binds in a sequence specific manner to the consensus sequence 5'(T/A)N(T/C)N(G/A)NC(T/A)TCC(T/A)N(T/A)(T/A)(T/G)-3' with a high affinity (K_d = 2.9 x 10^{-10}M), but low specificity. The DNA is contorted when bound in the Gcr1p-DNA nucleoprotein complex. The consensus sequence derived from the study of Gcr1p DNA-binding corresponds to known *cis*-element demonstrated by genetic methods to be important for *GCR1* dependent transcription. The results are consistent with the proposal that additional proteins are important for facilitation of Gcr1p binding and activation at the glycolytic gene upstream activating sequences.

REFERENCES

Affolter, M., Percival-Smith, A., Muller, M., and Leupin W. (1990). DNA binding properties of the purified *Antennapedia* homeodomain. Poc. Natl. Acad. sci. USA *87*, 4093-4097.

Allison, L.A., Moyle, M., Shales, M., and Ingles, C.J. (1985). Extensive homology among the largest subunits of eukaryotic and prokaryotic RNA polymerases. Cell *42*, 599-610.

Allison, L.A., Wong, J.K-C., Fitzpatrick, V.D., Moyle, M., and Ingles, C.J. (1988). The C-terminal domain of the largest subunit of RNA polymerase II of *Saccharomyces cerevisiae*, *Drosophila melanogaster*, and mammals: a conserved structure with an essential function. Mol. Cell. Biol. *8*, 321-329.

Amann, E., Brosius, J., and Ptashne, M. (1983). Vectors bearing a hybrid *trp-lac* promoter useful for regulated expression of cloned genes in *Escherichia coli*. Gene *25*, 167-178.

Amati, B.B. and Gasser, S.M. (1988). Chromosomal ARS and CEN elements bind specifically to the yeast nuclear scaffold. Cell *54*, 967-978.

Anderson, J.N. (1986). Detection, sequence patterns and function of unusual DNA structures. Nucleic Acids Res. *14*, 8513-8533.

Arndt, K.T., Styles, C.A., and Fink, G.R. (1989). A suppressor of a *HIS4* transcriptional defect encodes a protein with homology to the catalytic subunit of protein phosphatase. Cell *56*, 527-537.

Aso, T., Vasavada, H.A., Kawaguchi, T., Germino, F.J., Ganguly, S., Kitajima, s., Weissman, S.M., and Yasukochi, Y. (1992). Characterization of cDNA for the large subunit of the transcription initiation factor TFIIF. Nature *355*, 461-464.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K. (1989). Current Protocols in Molecular Biology (New York, New York: Greene Publishing Associates and Wiley-Interscience).

Baker, H.V. (1986). Glycolytic gene expression in *Saccharomyces cerevisiae*: nucleotide sequence of *GCR1*, null mutations, and evidence for expression. Mol. Cell. Biol. *6*, 3774-3784.

Baker, H.V. (1991). *GCR1* of *Saccharomyces cerevisiae* encodes a DNA binding protein whose binding is abolished by mutations in the CTTCC sequence motif. Poc. Natl. Acad. sci. USA *88*, 9443-9447.

Bennetzen, J.L. and Hall, B.D. (1982). Codon selection in yeast. J. Biol. Chem. *257*, 3026-3031.

Berger, S.L., Cress, W.D., Cress, A., Triezenberg, S.J., and Guarente, L. (1990). Selective inhibition of activated but not basal transcription by the acidic activation domain of VP16: evidence for transcriptional adaptors. Cell *61*, 1199-1208.

Berger, S., Pina, B., Silverman, G., Marcus, G., Agapite, J., Regier, J., Treizenberg, S., and Guarente, L. (1992). Genetic isolation of ADA2: a potential transcriptional adaptor required for function of certain acidic activation domains. Cell *70*, 251-266.

Biednenkapp, H., Borgmeyer, U., Sippel, A.E., and Klempnauer, K-H. (1988). Viral *myb* oncogene encodes a sequence-specific DNA-binding activity. Nature *335*, 835-837.

Bitter, G.A., Chang, K.K.H., and Egan, K.M. (1991). A multi-component upstream activation sequence of the *Saccharomyces cerevisiae* glyceraldehyde-3-phosphate dehydrogenase gene promoter. Mol. Gen. Genet. *231*, 22-32.

Bitter, G.A. and Egan, K. (1984). Expression of heterologous genes in *Saccharomyces cerevisiae* from vectors utilizing the glyceraldehyde-3-phosphate dehydrogenase gene promoter. Gene *32*, 263-274.

Bitter, G.A. and Egan, K. (1988). Expression of heterologous genes in *S. cerevisiae* from vectors utilizing the glyceraldehyde-3-phosphate dehydrogenase gene promoter. Gene *69*,193-207.

Bitter, G.A., Egan, K.M., Koski, R.A., Jones, M.O., Elliott, S.G., and Giffin J.C. (1987). Expression and secretion vectors for yeast. Methods Enzymol *153*, 516-544.

Bonnefoy, E. and Rouviere-Yaniv, J. (1991). HU and IHF, two homologous histone-like proteins of Escherichia coli, form different protein-DNA complexes with short DNA fragments. EMBO J. *10*, 687-696.

Bracco, L., Kotlarz, D., Kolb, A., Diekmann, S., and Buc, H. (1989). Synthetic curved DNA sequences can act as transcriptional activators in *Escherichia coli*. EMBO J. *8*, 4289-4296.

Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. *72*, 248-254.

Brandl, C.J. and Struhl, K. (1990). A nucleosome-positioning sequence is required for GCN4 to activate transcription in the absence of a TATA element. Mol. Cell. Biol. *10*, 4256-4265.

Brent, R. and Ptashne, M. (1985). A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. Cell *43*, 729-736.

Brindle, P.K., Holland, J.P., Willett, C.E., Innis, M.A., and Holland, M.J. (1990). Multiple factors bind the upstream activation sites of the yeast enolase genes *ENO1* and *ENO2*: ABF1 protein, like repressor activator protein RAP1, binds *cis*-acting sequences which modulate repression or activation of transcription. Mol. Cell. Biol. *10*, 4872-4895.

Britten, R.J. and Davidson, E.H. (1969). Gene regulation for higher cells: a theory. Science *165*, 349-358.

Bucher, P. (1990). Weight matrix descriptions of four eukaryotic RNA polymerase II promoter elements derived from 502 unrelated promoter sequences. J. Mol. Biol. *212*, 563-578.

Buchman, A., Kimmerley, W.J., Rine, J., and Kornberg, R.D. (1988). Two DNA-binding factors recognize specific sequences at silencers, upstream activating sequences, autonomously replicating sequences, and telomeres in *Saccharomyces cerevisiae*. Mol. Cell. Biol. *8*, 210-225.

Buchman, A., Lue, N.F., and Kornberg, R.D. (1988). Connections between transcriptional activators, silencers, and telomeres as revealed by functional analysis of a yeast DNA-binding protein. Mol. Cell. Biol. *8*, 5086-5099.

Buratowski, S., Hahn, S., Sharp, P.A., and Guarente, L. (1988). Function of a yeast TATA element-binding protein in a mammalian transcription system. Nature *334*, 37-42.

Butler, G., Dawes, I.W., and McConnell, D.J. (1990). TUF factor binds to the upstream region of the pyruvate decarboxylase structural gene (*PDC1*) of *Saccharomyece cerevisiae*. Mol. Gen. Genet. *223*, 449-456.

Cagniard-Latour (1938). Memoire sur la fermentation vineuse. Ann. Chim *68*, 206-223.

Calladine, C.R., Collis, C.M., Drew, H.R., and Mott, M.R. (1991). A study of electrophoretic mobility of DNA in agarose and polyacrylamide gels. J. Mol. Biol. *221*, 981-1005.

Capieux, E., Vignais, M.L., Sentenac, A., and Goffeau, A. (1989). The yeast H^+-ATPase gene is controlled by the promoter binding factor TUF. J. Biol. Chem. *264*, 7437-7446.

Cardenas, M.E., LaRoche, T.L., and Gasser, S.M. (1990). The composition and morphology of the yeast nuclear scaffolds. J. Cell Sci. *96*, 439-450.

Carey, J. (1991). Gel retardation. Meth. Enzymol. *208*, 103-117.

Casadaban, M.J. and Cohen, S.N. (1980). Analysis of gene control signals by DNA fusion and coloning in *Escherichia coli*. J. Mol. Biol. *138*, 179-207.

Cavallini, B., Huet, J., Plassat, J-L., Sentenac, A., Egly, J-M., and Chambon, P. (1988). A yeast activity can substitute for the HeLa cell TATA box factor. Nature *334*, 77-80.

Chadwick, P., Pirrotta, V., Steinberg, R., Hopkins, N., and Ptashne, M. (1970). The lambda and 434 phage repressors. Cold Spring Harbor Symposia on Quantitative Biology *XXXV*, 283-294.

Chambers, A., Stanway, C., Kingsman, A.J., and Kingsman, S.M. (1988). The UAS of the yeast *PGK* gene is composed of multiple functional elements. Nucleic Acids Res. *16*, 8245-8260.

Chambers, A., Stanway, C., Tsang, J.S.H., Henry, Y., Kingsman, A.J., and Kingsman, S.M. (1990). ARS binding factor 1 binds adjacent to RAP1 at the UASs of the yeast glycolytic genes *PGK* and *PYK*. Nucleic Acids Res. *18*, 5393-5399.

Chambers, A., Tsang, J.S.H., Stanway, C., Kingsman, A.J., and Kingsman, S.M. (1989). Transcriptional control of the *Saccharomyces cerevisiae PGK* gene by RAP1. Mol. Cell. Biol. *9*, 5516-5524.

Chasman, D.I., Lue, N.F., Buchman, A., LaPointe, J.W., Lorch, Y., and Kornberg, R.D. (1990). A yeast protein that influences the chromatin structure of UAS_G and functions as a powerful auxiliary gene activator. Genes Dev. *4*, 503-514.

Chittenden, T., Livingston, D.M., and Kaelin Jr, W.G. (1991). The T/E1A-binding domain of the Retinoblastoma product can interact selectively with a sequence-specific DNA-binding protein. Cell *65*, 1073-1082.

Clifton, D. and Fraenkel, D.G. (1981). The *gcr1* (glycolysis regulation) mutation of *Saccharomyces cerevisiae*. J. Biol. Chem. *256*, 13074-13078.

Clifton, D., Weinstock, D., and Fraenkel, D.G. (1978). Glycolysis mutants of *Saccharomyces cerevisiae*. Genetics *88*, 1-11.

Cohen, R., Holland, J.P., Yokoi, T., and Holland, M.J. (1986). Identification of a regulatory region that mediates glucose-dependent induction of the *Saccharomyces cerevisiae* enolase gene *ENO2* Mol. Cell. Biol. *6*, 2287-2297.

Cohen, R., Yokoi, T., Holland, J.P., Pepper, A.E., and Holland, M.J. (1987). Transcription of the constitutively expressed yeast enolase gene *ENO1* is mediated by positive and negative *cis*-acting regulatory sequences. Mol. Cell. Biol. *7*, 2753-2761.

Colgan, J. and Manley, J.L. (1992). TFIID can be rate limiting *in vivo* for TATA-containing, but not TATA-lacking, RNA polymerase II promoters. Genes Dev. *6*, 304-315.

Comai, L., tanese, N. and Tjian, R. (1992). The TATA-binding protein and associated factors are integral components of the RNA polymerase I transcription factor, SL1. Cell *68*, 965-976.

Conrad, M.N., Wright, J.H., Wolf, A.J., and Zakian, V.A. (1990). RAP1 protein interacts with yeast telomeres *in vivo*: overproduction alters telomere structure and decreases chromosome stability. Cell *63*, 739-750.

Corden, J.L., Caden, D.L., Ahearn Jr, J.M., and Dahmus, M.E. (1985). A unique structure at the carboxyl terminus of the largest subunit of eukaryotic RNA polymerase II. Poc. Natl. Acad. Sci. USA *82*, 7934-7938.

Corsetti, M. Briata, P., Sanseverino, L. Daga, A., Airoldi, I., Simeon, A., Palmisano, G., Angelini, C., Boncinelli, E., and Corte, G. (1992). Differential DNA binding protperties of three human homeodomain proteins. Nucleic Acids Res. *20*, 4465-4472.

Cortes, P., Flores, O., and Reinberg, D. (1992). Factors involved in specific transcription by mammalian RNA polymerase II: purification and analysis of transcription factor IIA and identification of transcription factor IIJ. Mol. Cell. Biol. *12*, 413-421.

Courey, A. and Tjian, R. (1988). Analysis of Sp1 *in vivo* reveals multiple transcriptional domains including a novel glutamine-rich activation motif. Cell *55*, 887-898.

Craig, N. and Nash, H. (1984). *E. coli* integration host factor binds to specific sites in DNA. Cell *39*, 707-716.

Cress, A, and Triezenberg, S. (1991). Critical structural elements of the VP16 transriptional activation domain. Science *251*, 87-90.

Crothers, D.M., Gartenberg, M.R., and Shrader, T.E. (1991). DNA bending in protein-DNA complexes. Meth. Enzymol. *208*, 118-146.

Davison, B.L., Egly, J-M., Mulvihill, E.R., and Chambon, P. (1983). Formation of stable preinitiation complexes between eukaryotic class B transcription factors and promoter sequences. Nature *301*, 680-686.

Delbrück, M. (1975). Mind from matter? An essay on evolutionary epistemology (Carlton, Victoria: Blackwell Scientific Publications).

Desplan, C., Theis, J., and O'Farrell, P.H. (1985). The *Drosophila* development gene, *engrailed*, encodes a sequence-specific DNA binding activity. Science *318*, 630-635.

Devlin, C., Tice-Baldwin, K., Shore, D., and Arndt, K.T. (1991). RAP1 is required for BAS1/BAS2- and GCN4-dependent transcription of the yeast *HIS4* gene. Mol. Cell. Biol. *11*, 3642-3651.

Diffley, J.F.X. and Stillman, B. (1989). Similarity between the transcriptional silencer binding proteins ABF1 and RAP1. Science *246*, 1034-1038.

Drazinic, C.M., Lopez, M.C., and Baker H.V. (1994). Native Gcr1p requires a functional CT-box to mediate *GCR1*-dependent gene expression. Submitted.

Drew, H.R. and Travers, A.A. (1985). Structural junctions in DNA: the influence of flanking sequence on nuclease digestion specificities. Nucleic Acids Res. *13*, 4445-4467.

Dynan, W.S. (1989). Modularity in promoters and enhancers. Cell *58*, 1-4.

Dynlacht, B.D., Hoey, T., and Tjian, R. (1991). Isolation of coactivators associated with the TATA-binding protein that mediate transcriptional activation. Cell *66*, 563-576.

Eckdahl, T.T. and Anderson, J.N. (1990). Conserved DNA structures in origins or replication. Nucleic Acids Res. *18*, 1609-1612.

Eisenmann, D.M., Dollard, C., and Winston, F. (1989). *SPT15*, the gene encoding the yeast TATA binding factor TFIID, is required for normal transcription initiation *in vivo*. Cell *58*, 1183-1191.

Fassler, J.S. and Winston, F. (1989). The *Saccharomyces cerevisiae SPT13/GAL11* gene has both positive and negative regulatory roles in transcription. Mol. Cell. Biol. *9*, 5602-5609.

Fedor, M.J., Lue, N.F., and Kornberg, R.D. (1988). Statistical positioning of nucleosomes by specific protein-binding to an upstream activating sequence in yeast. J. Mol. Biol. *204*, 109-127.

Felsenfeld, G. (1992). Chromatin as an essential part of the transcriptional mechanism. Nature *355*, 219-224.

Ferrari, S., Harley, V.R., Pontiggia, A., Goodfellow, P.N., Lovell-Badge, R., and Bianchi, M. (1992). SRY, like HMG1, recognizes sharp angles in DNA. EMBO J. *11*, 4497-4506.

Fields, S. and Song, O-K. (1989). A novel genetic system to detect protein-protein interactions. Nature *340*, 245-246.
Fikes, J.D., Becker, D.M., Winston, F., and Guarente, L. (1990). Striking conservation of TFIID in *Schizosaccharomyces pombe* and *Saccharomyces cerevisiae*. Nature *346*, 291-294.

Finkelstein, A., Kostrab, C.F., Li, J., Chavez, D.P., Wang, B.Q., Fang, S.M., Greenblatt, J., and Burton, Z.F. (1992). A cDNA encoding RAP74, a general initiation factor for transcription by RNA polymerase II. Nature *355*, 464-467.

Fitzpatrick, V., Percival-Smith, A., Ingles, C., and Krause, H. (1992). Homeodomain-independent activity of the fushi tarazu polypeptide in Drosophila embryos. Nature *356*, 610-612.

Flanagan, P.M., Kelleher III, R.J., Sayre, M.H., Tschochner, H., and Kornberg, R.D. (1991). A mediator required for activation of RNA polymerase II transcription *in vitro*. Nature *350*, 436-438.

Flashner, Y. and Gralla, J. (1988). DNA dynamic flexibility and protein recognition: differential stimulation by bacterial histone-like protein HU. Cell *54*, 713-721.

Florence, B., Handrow, R., and Laughon, A. (1991). DNA-binding specificity of the *fushi tarazu* homeodomain. Mol. Cell. Biol. *11*, 3613-3623.

Fraenkel, D.G. (1982). Carbohydrate metabolism. In The Molecular Biology of the yeast *Saccharomyces cerevisiae*: Metabolism and Gene Expression. J.N. Strathern, E.W. Jones, and J.R. Borach, eds. (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press), pp. 1-37.

Frankel, A.D. and Kim, P.S. (1991). Modular structure of transcription factors: implications for gene regulation. Cell *65*, 717-719.

Fried, M. and Crothers, D.M. (1981). Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. Nucleic Acids Res. *9*,6505-6525.

Fruton, J.S. (1972). Molecules and Life. (Wiley, New Yorkl).

Fruton, J.s. (1976). The emergence of biochemistry. Science *192*, 327-334.

Ganster R.W., Shen W., and Schmidt, M.C. (1991). Isolation of STD1, a high-copy-number suppressor of a dominant negative mutation in the yeast TATA-binding protein. Mol. Cell. Biol. *13*,3650-3659.

Garner, M.M. and Revzin, A. (1981). A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: applications to components of *Escherichia colii* lactose operon regulatory system. Nucleic Acids Res. *9*, 3037-3060.

Gasch, A., Hoffmann, A., Horikoshi, M., Roeder, R.G., and Chua, N-H. (1990). *Arabidopsis thaliana* contains two genes for TFIID. Nature *346*, 390-394.

Gasser, S.M. and Laemmli, U.K. (1986). Cohabitation of acaffold binding regions with upstream/enhancer elements of three developmentally regulated genes. Cell *46*, 521-530.

Gasser, S.M. and Laemmli, U.K. (1987). A glimpse at chromosomal order. Trends Gens. *3*, 16-22.

Gay-Lassic, J.L. (1810). Extrait d'un memoire surla fermentation. Ann. Chim *76*, 245-259.

Giese, K., Amsterdam, A., and Gosschedl, R. (1991). DNA-binding properties of the HMG domain of the lymphoid-specific transcriptional regulator LEF-1. Genes Dev. *5*,2567-2578.

Gill, G., Pascal, E., Tseng, Z., and Tjian, R. (1994). A glutamine-rich hydrophobic patch in transcription factor Sp1 contacts the dTAFII10 component of the Drosophila TFIID complex and mediates transcriptional activation. Proc. Natl. Acad. Sci. USA *91*, 192-196.

Gill, G. and Ptashne, M. (1988). Negative effect of the transcriptional activator *GAL4*. Nature *334*, 721-724.

Goodman, S.D. and Nash, H.A. (1989). Functional replacement of a protein-induced bend in a DNA recombination site. Nature *341*, 251-254.

Goodrich, J., Hoey, T., Thut, C., Admon, A., and Tjian, R. (1993). *Drosophila* TAFII40 interacts with both a VP16 activation domain and the basal transcription factor TFIIB. Cell *75*, 519-530.

Gould, S.J. (1977). Ontogeny and phylogeny (Cambridge, Massachusetts: Harvard University Press).

Green, S. and Chambon, P. (1987). Oestradiol induction of a glucocorticoid-responsive gene by a chimaeric receptor. Nature *325*, 75-78.

Greenblatt, J. (1991a). Roles of TFIID in transcriptional initiation by RNA polymerase II. Cell *66*, 1067-1070.

Greenblatt, J. (1991b). RNA polymerase-associated transcription factors. Trends Gens. *16*, 408-411.

Guarente, L. (1988). UASs and enhancers: common mechanism of transcriptional activation in yeast and mammals. Cell *52*, 303-305.

Guarente, L. and Bermingham-McDonogh, O. (1992). Conservation and evolution of transcriptional mechanism in eukaryotes. Trends Gens. *8*, 27-32.

Haass, M.M. and Feix, G. (1992). Two different cDNAs encoding TFIID proteins of maize. FEBS Lett. *301*, 294-298.
Hahn, S., Buratowski, S., Sharp, P.A., and Guarente, L. (1989). Isolation of the gene encoding the yeast TATA binding protein TFIID: a gene identical to the *SPT15* suppressor of Ty element insertions. Cell *58*, 1173-1181.

Hahn, S. (1993). Efficiency in activation. Science *363*, 672-673.

Hanahan, D. (1983). Studies on transformation of *Escherichia colia* with plasmids. J. Mol. Biol. *166*, 557-580.

Hardy, C.F.J., Balderes, D., and Shore, D. (1992). Dissection of a carboxy-terminal region of the yeast regulatory protein RAP1 with effects on both transcriptional activation and silencing. Mol. Cell. Biol. *12*, 1209-1217.

Harley, V., Jackson, D., Hextall, P., Hawkins, J., Berkovitz, G., Sockanathan, S., Lovell-Badge, R., and Goodfellow, P.N. (1992). DNA binding activity of recombinant SRY from normal males and XY females. Science *255*, 453-456.

Harrison, S. and Aggarwal, A. (1990). DNA recognition by proteins with the helix-turn-helix motif. Annu. Rev. Biochem. *59*, 933-969.

Hawley, D. (1991). Transcriptional activation: enter TFIIB. Trends Biol. Sci. *16*, 317-318.

Heisenberg, W. (1971). Physics and Beyond. (New York: Harper and Row Publishers).

Henry, Y.A., Chambers, A., Tsang, J.S.H., Kingsman, A.J., and Kingsman, S.M. (1990). Characterization of the DNA binding domain of the yeast RAP1 protein. Nucleic Acids Res. *18*, 2617-2623.

Hess, B., Bioteux, A., and Kruger, J. (1969). Cooperation of glycolytic enzymes. Adv. Enzyme Regl. *7*, 149-169.

Himmelfarb, H.J., Pearlberg, J., Last, D.H., and Ptashne, M. (1990). GAL11P: a yeast mutation that potentiates the effect of weak GAL4-derived activators. Cell *63*, 1299-1309.

Hochschild, A. and Ptashne, M. (1986). Cooperative binding of lambda repressors to sites separated by intergral turns of the DNA helix. Cell *44*, 681-687.

Hoey, T., Dynlacht, B.D., Peterson, M.G., Pugh, B.F., and Tjian, R. (1990). Isolation and characterization of the *Drosophila* gene encoding the TATA box binding protein, TFIID. Cell *61*, 1179-1186.

Hoffmann, A., Sinn, E., Yamamoto, T., Wang, J., Roy, A., Horikoshi, M., and Roeder, R.G. (1990). Highly conserved core domain and unique N terminus with presumptive regulatory motifs in a human TATA factor (TFIID). Nature *346*, 387-390.

Hofmann, J.F-X., Laroche, T., Brand, A.H., and Gasser, S.M. (1989). RAP-1 loop formation is necessary for DNA loop formation *in vitro* at the silent mating type locus *HML*. Cell *57*, 725-737.

Holland, M.J., and Holland, J.P. (1978). Isolation and identification of yeast messenger ribonucleic acids coding for enolase, glyceraldehyde-3-phosphate dehydrogenase, and phosphoglycerate kinase. Biochemistry *17*, 4900-4907.

Holland, M.J., Yokoi, T., Holland, J.P., Myambo, K., and Innis, M.A. (1987). The *GCR1* gene encodes a positive transcriptional regulator of the enolase and glyceraldehyde-3-phosphate dehydrogenase gene families in *Saccharomyces cerevisiae*. Mol. Cell. Biol. *7*, 813-820.

Hollenberg, S.M. and Evans, R.M. (1988). Multiple and cooperative *Trans*-activation domains of the human glucocorticoid receptor. Cell *55*, 899-906.

Hope, I.A. and Struhl, K. (1987). GCN4, eukaryotic transcriptional activator protein, binds as a diamer to target DNA. EMBO J. *6*, 2781-2784.

Horikoshi, M., Bertuccioli, C., Takada, R., Wang, J., Yamamoto, T., and Roeder, R.G. (1992). Transcription factor TFIID induces DNA bending upon binding to the TATA element. Poc. Natl. Acad. sci. USA *89*, 1060-1064.

Horikoshi, M., Wang, C.K., Fujii, H., Cromlish, J.A., Weil, P.A., and Roeder, R.G. (1989). Cloning and structure of a yeast gene encoding a general transcription initiation factor TFIID that binds to the TATA box. Nature *341*, 299-303.

Huet, J., Cottrelle, P., Cool, M., Vignais, M.L., Thiele, D., Marck, C., Buhler, J-M., Sentenac, A., and Fromageot, P. (1985). A general upstream binding factor for genes of the yeast translational apparatus. EMBO J. *4*, 3539-3547.

Huie, M.A., Scott, E.W., Drazinic, C.M., Lopez, M.C., Hornstra, I.K., Yang, T.P., and Baker, H.V. (1992). Characterization of the DNA-binding activity of GCR1: *in vivo* evidence for two GCR1-binding sites in the upstream activating sequence of *TPI* of *Saccharomyces cerevisiae*. Mol. Cell. Biol. *12*, 2690-2700.

Irvine, D., Tuerk, C., and Gold, L. (1991). Systematic evolution of ligand by exponential enrichment with integrated optimization by non-linear analysis. J. Mol. Biol. *222*, 739-761.

Irwin, N. and Ptashne, M. (1987). Mutants of the catabolite activator protein of *Escherichia coli* that are specifically deficient in the gene-activation function. Poc. Natl. Acad. sci. USA *84*, 8315-8319.

Jacob, F. (1977). Evolution and tinkering. Science *196*, 1161.

Janknecht, R., Hipskind, R., Houthaeve, T., Nordheim, A., and Stunnenberg, H. (1992). Identification of multiple SRF N-terminal phosphoryation sites affecting DNA binding properties. EMBO J. *11*,1045-1054.

Janson, L. and Pettersson, U. (1990). Cooperative interactions between transcription factors Sp1 and OTF-1. Proc. Natl. Acad. Sci. USA *87*, 4732-4736.

Jantzen, H., Chow, A., King, D. and Tjian, R. (1992). Multiple domains of the RNA polymerase I activator hUBF interact with the TATA-binding protein complex hSL1 to mediate transcription. Genes Dev. *6*, 1950-1963.

Jordan, S. and Pabo, C.O. (1988). Structure of the lambda complex at 2.5A resolution: details of the repressor-operator interactions. Science *242*, 893-907.

Ju, Q., Morrow, B.E., and Warner, J.R. (1990). REB1, a yeast DNA-binding protein with many targets, is essential for growth and bears some resemblance to the oncogene *myb*. Mol. Cell. Biol. *10*, 5226-5234.

Kahn, J. and Crothers, D. (1992). Protein-induced bending and DNA cyclization. Proc. Natl. Acad. Sci. USA *89*, 6343-6347.

Kahn, J. and Crothers, D. (1994). Cetection of localized DNA flexibility. Nature *368*, 163-166.

Kakidani, H. and Ptashne, M. (1987). GAL4 activates gene expression in mammalian cells. Cell *52*, 161-167.

Kao, C.C., Lieberman, P.M., Schmidt, M.C., Zhou, Q., Pei, R., and Berk, A.J. (1990). Cloning of a transcriptionally active human TATA binding factor. Science *248*, 1646-1650.

Karim, F.D., Urness, L.D., Thummel, C.S., Klemsz, M.J., McKercher, S.R., Celada, A., VanBeveren, C., Maki, R.A., Gunther, C.V., Nye, J.A., and Graves, B.J. (1990). The ETS-domain: a new DNA-binding motif that recognizes a purine-rich core DNA sequence. Genes Dev. *4*, 1451-1453.

Karlin, S. and Brendel, V. (1992). Chance and statistical significance in protein DNA sequence analysis. Science *257*, 39-49.

Kassavetis, G.A., Bartholomew, B., Blanco, J.A., Johnson, T.E., and Geiduschek, E.P. (1991). Two essential components of the *Saccharomyces cerevisiae* transcription factor TFIIB: transcription and DNA-binding properties. Poc. Natl. Acad. sci. USA *88*, 7308-7312.

Kawasaki, G. and Fraenkel, D.G. (1982). Cloning of yeast glycolysis genes by complementation. Biochem. Biophys. Res. Commun. *108*, 1107-1112.

Kelleher III, R.J., Flanagan, P.M., Chasman, D.I., Ponticelli, A.S, Struhl, K., and Kornberg, R.D. (1992). Yeast and human TFIIDs are interchangeable for the response to acidic transcriptional activators *in vitro*. Genes Dev. *6*, 296-303.

Kerppola, T.K. and Curran, T. (1991). DNA bending by Fos and Jun: the flexible hinge model. Science *254*, 1210-1214.

Kierkegaard, S. (1975). Journals and Papers. (translated by H.V. Hong and E.H. Hong) Bloomington: Indiana University Press.

Kim J., Nikolov, D. and Burley, S. (1993). Co-crystal structure of TBP recognizing the minor groove of a TATA element. Nature *365*, 520-527.

Kim, J., Zwieb, C., Wu, C., and Adhy, S. (1989). Bending of DNA by gene-regulatory proteins: construction and use of a DNA bending vector. Gene *85*, 15-23.

Kimmerley, W.J., Buchman, A., Kornberg, R.D., and Rine, J. (1988). Roles of two DNA-binding factors in replication, segregation and transcriptional repression mediated by a yeast silencer. EMBO J. *7*, 2241-2253.

King, M.C. and Wilson, A.C. (1975). Evolution on two levels in human and chimpanzee. Science *188*, 107-116.

Kohler, R.E. (1972). The Enzyme Theory and the Origin of Biochemistry. (Philadelphia, W.B. Saunders).

Kornberg, A. (1989). For the Love of Enzymes. (Cambridge, Mass).

Kornberg, R.D. and Lorch, Y. (1991). Irresistible force meets immovable object: transcription and the nucleosome. Cell *67*, 833-836.

Lahm, A. and Suck, D. (1991). DNaseI-induced DNA conformation. 2A structure of a DNaseI-octamer complex. J. Mol. Biol. *222*, 645-667.

Lavoisier, A.I. (1789). Elements of Chemistry, translated by R. Kerr (Creech. Edinburgh).

Leidig, F., Shepard, A.R., Zhang, W., Stelter, A., Cattini, P.A., Baxter, J.D., and Eberhardt, N.L. (1992). Thyroid hormone responsiveness in human growth hormone-related genes. J. Biol. Chem. *267*, 913-921.

Letovsky, J. and Dynan, W.S. (1989). Measurement of the binding of transcription factor Sp1 to a single GC box recognition sequence. Nucleic Acids Res. *13*, 2639-2653.

Leuther, K.K., Salmeron, J.M., and Johnston, S.A. (1993). Genetic evidence that an activation domain of GAL4 does not require acidity and may form a β sheet. Cell *72*, 575-585.

Levine, A.S. (1989). A new reductionism and a new journal. New Biol. *1*, 1-2.

Lewin, B. (1990). Commitment and activation at pol II promoters: a tail of protein-protein interactions. Cell *61*, 1161-1164.

Lieberman, P.M., Schmidt, M.C., Kao, C.C., and Berk, A.J. (1991). Two distinct domains in the yeast transcription factor IID and evidence for a TATA box-induced conformational change. Mol. Cell. Biol. *11*, 63-74.

Lin, Y-S. and Green, M.R. (1991). Mechanism of action of an acidic transcriptional activation *vitro*. Cell *64*, 971-981.

Longtine, M.S., Wilson, N.M., Petracek, M.E., and Berman, J. (1989). A yeast telomere binding activity binds to two related telomere sequence motifs and is indistinguishable from RAP1. Curr. Genet. *16*, 225-239.

Luscher, B., Christenson, E., Litchfield, D.W., Krebs, E.G., and Eisenman, R.N. (1990). Myb DNA binding inhibited by phosphorylation at a site deleted during oncogenic activation. Science *344*, 517-522.

Lustig, A.J., Kurtz, S., and Shore, D. (1990). Involvement of the silencer and UAS binding protein RAP1 in regulation of telomere length. Science *250*, 549-553.

Ma, J. and Ptashne, M. (1987a). A new class of yeast transcriptional activators. Cell *51*, 113-119.

Ma, J. and Ptashne, M. (1987b). Deletion analysis of GAL4 defines two transcriptional activating segments. Cell *48*, 847-853.

Machida, M., Jigami, Y., and Tanaka, H. (1989). Purification and characterization of a nuclear factor which binds specifically to the upstream activation sequence of *Saccharomyces cerevisiae* enolase 1 gene. Eur. J. Biochem. *184*, 305-311.

Malik, S., Hisatake, K., Sumimoto, H., Horikoshi, M., and Roeder, R.G. (1991). Sequence of general transcription factor TFIIB and relationships to other initiation factors. Poc. Natl. Acad. sci. USA *88*, 9553-9557.

Marahrens, Y. and Stillman, B. (1992). A yeast chromosomal origin of DNA replication defined by multiple functional elements. Science *255*, 817-823.

Marshall, M., Mahoney, D., Rose, A., Hicks, J.B., and Broach, J.R. (1987). Functional domains of *SIR4*, a gene required for position effect regulation in *Saccharomyces cerevisiae*. Mol. Cell. Biol. *7*, 4441-4452.

Martin, K.J. (1991). The interactions of transcription factors and their adaptors, coactivators and accessory proteins. BioEssays *13*, 499-503.

Matsui, T., Segall, J., Weil, P.A., and Roeder, R.G. (1980). Multiple factors required for accurate initiation of transcription by purified RNA polymerase II. J. Biol. Chem. *255*, 11992-11996.

Mayr, E. (1970). Population, Species, and Evolution (Cambridge, Massachusetts: Harvard University Press).

McNeil, J.B., Dykshoorn, P., Huy, J.N., and Small, S. (1990). The DNA-binding protein RAP1 is required for efficient transcriptional activation of the yeast *PYK* glycolytic gene. Curr. Genet. *18*, 405-412.

Mermod, N., O'Neill, E.A., Kelly, T.J., and Tjian, R. (1989). The proline-rich transcriptional activator of CTF/NF-1 is distinct from the replication and DNA binding domain. Cell *58*, 741-753.

Messing, J. New M13 vectors for cloning. (1983). Methods Enzymol. *101*, 20-78.

Meyers, R. and Sharp, P. (1993). TATA-binding protein and asociated factors in polymerase II and polymerase III transcription. Mol. Cell. Biol. *13*, 7953-7960.

Mitchell, P.J. and Tjian, R. (1989). Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. Science *245*, 371-378.

Moehle, C.M. and Hinnebusch, A.G. (1991). Association of RAP1 binding sites with stringent control of ribosomal protein gene transcription in *Saccharomyces cerevisiae*. Mol. Cell. Biol. *11*, 2723-2735.

Morrow, B.E., Johnson, S.P., and Warner, J.R. (1989). Proteins that bind to the yeast rDNA enhancer. J. Biol. Chem. *264*, 9061-9068.

Morrow, B.E., Ju, Q., and Warner, J.R. (1990). Purification and characterization of the yeast rDNA binding protein REB1. J. Biol. Chem. *265*, 20778-20783.

Morse, R.H. (1992). Transcribed chromatin. Trends Biol. Sci. *17*, 23-26.

Muhich, M.L., Iida, C.T., Horikoshi, M. Roeder, R. and Parker, C.S. (1990). cDNA clone encoding Drosophila transcription factor TFIID. Poc. Natl. Acad. sci. USA *87*,9148-9152.

Nasmyth, K.A. (1982). The regulation of yeast mating-type chromatin structure by *SIR*: an action at a distance affecting both transcription and transposition. Cell *30*, 567-578.

Nelson, H.B. and Laughon, A. (1990). The DNA binding specificity of the *Drosophila fushi tarazu* protein: a possible role for DNA bending in homeodomain recognition. New Biol. *2*, 171-178.

Newton, I. (1718). Opticks: or a Treatis of the Reflections, Refractions, Inflections and Colours of Light. (Innys, London).

Nietzsche, F. (1967). The Will to Power. (translated by W. Kaufmann) New York: Random House.

Nikolov, D., Hu, S. Lin, J. Gasch, A., Hoffmann, A. Horikoshi, M., Chua, N., Roeder, R., and Burley, S. (1992). Crystal structure of TFIID TATA-box binding protein. Nature *360*, 40-46.

Nishizawa, M., Araki, R., and Teranishi, Y. (1989). Identification of an upstream activating sequence and an upstream repressible sequence of the pyruvate kinase gene of the yeast *Saccharomyces cerevisiae*. Mol. Cell. Biol. *9*, 442-451.

Nishizawa, M., Suzuki, Y., Nogi, Y., Matsumoto, K., and Fukasawa, T. (1990). Yeast Gal11 protein mediates the transcriptional activation signal of two different transacting factors, Gal4 and general regulatory factor I/repressor/activator site binding protein 1/translation upstream factor. Poc. Natl. Acad. sci. USA *87*, 5373-5377.

Ogden, J.E., Stanway, C., Kim, S., Mellor, J., Kingsman, A.J., and Kingsman, S.M. (1986). Efficient expression of the *Saccharomyces cerevisiae PGK* gene depends on an upstream activating sequence but does not require TATA sequences. Mol. Cell. Biol. *6*, 4334-4343.

Ohkuma, Y., Sumimoto, H., Hoffmann, A., Shimasaki, S., Horikoshi, M., and Roeder, R.G. (1991). Structural motifs and potential homologies in the large subunit of human general transcription factor TFIIE. Nature *354*, 398-401.

Oliphant, A.R. and Struhl, K. (1987). The use of random-sequence oligonucleotides for determining consensus sequences. Meth. Enzymol. *155*, 568-582.

Oliphant, A.R. and Struhl, K. (1988). Defining the consensus sequence of *E. coli* promoter elements by random selection. Nucleic Acids Res. *16*, 7673-7683.

Oliphant, A.R., Brandl, C.J., and Struhl, K. (1989). Defining the sequence specificity of DNA-binding proteins by selecting binding sites from random-sequence oligonucleotides: analysis of yeast GCN4 protein. Mol. Cell. Biol. *9*, 2944-2949.

Oliver, S.G. (1991). Classical yeast biotechnology. Biotechnology Handbooks *4*, 213-243.

O'Neil, K.T., Hoess, R.H., and DeGrado, W.F. (1990). Design of DNA-binding peptides based on the leucine zipper motif. Science *249*, 774-778.

Ouellette, M., Chen, J., wright, W., and Shay, J. (1992). Complexes containing the retinoblastoma gne product recognize different DNA motifs related to the E2F binding site. Oncogene *7*, 1075-1081.

Pavlovic, B. and Horz, W. (1988). The chromatin structure at the promoter of a glyceraldehyde phosphate dehydrogenase gene from *Saccharomyces cerevisiae* reflects its functional state. Mol. Cell. Biol. *8*, 5513-5520.

Pearson, W.R. (1990). Rapid and sensitive sequence comparison with FASTP and FASTA. Meth. Enzymol. *183*, 63-98.

Pearson, W.R. (1991). Searching protein sequence libraries: comparison of the sensitivity and selectivity of the Smith-Waterman and FASTA algorithms. Genomics *11*, 635-650.

Pearson, W.R. and Lipman, D.J. (1988). improved tools for biological sequence comparison. Poc. Natl. Acad. sci. USA *85*, 2444-2448.

Pérez-Martín, J. and Espinosa, M. (1991). The RepA repressor can act as a transcriptional activator by inducing DNA bends. EMBO J. *10*, 1375-1382.

Peterson, M.G., Inostroza, J., Maxon, M.E., Flores, O., Admon, A., Reinberg, D., and Tjian, R. (1991). Structure and functional properties of human general transcription factor IIE. Nature *354*, 369-373.

Peterson, M.G., Tanese, N., Pugh, B.F., and Tjian, R. (1990). Functional domains and upstream activation properties of cloned human TATA binding protein. Science *248*, 1625-1630.

Pfeifer, K., Arcangioli, B., and Guarente, L. (1987). Yeast HAP1 activator competes with the factor RC2 for binding to the upstream activation site UAS1 of the *CYC1* gene. Cell *49*, 9-18.

Polaczek, P. (1990). Bending of the origin of replication of *E. coli* by binding of IHF at a specific site. New Biol. *2*, 265-271.

Pollock, R. and Treisman, R. (1990). A sensitive method for the determination of protein-DNA binding specificities. Nucleic Acids Res. *18*, 6197-6204.

Prentki, P., Chandler, M., and Galas, D.J. (1987). *Escherichia coli* integration host factor bends the DNA at the ends of *IS1* and in an insertion hotspot with multiple IHF binding sites. EMBO J. *6*, 2479-2487.

Prentki, P., Pham, M-H., and Galas, D.J. (1987). Plasmid permutation vectors to monitor DNA bending. Nucleic Acids Res. *15*, 10060.

Ptashne, M. (1988). How eukaryotic transcriptional activators work. Nature *335*, 683-689.

Ptashne, M. and Gann, A.F. (1990). Activators and targets. Nature *346*, 329-331.

Pugh, B.F. and Tjian, R. (1990). Mechanism of transcriptional activation by Sp1: evidence for coactivators. Cell *61*, 1187-1197.

Pugh, B.F. and Tjian, R. (1992). Diverse transcriptional functions of the multisubunit eukaryotic TFIID complex. J. Biol. Chem. *267*, 679-682.

Reinberg, D., Horikoshi, M., and Roeder, R.G. (1987). Factors involved in specific transcription in mammalian RNA polymerase II. functional analysis of initiation factors IIA and identification of a new factor operating at sequences downstream of the initiation site. J. Biol. Chem. *262*, 3322-3330.

Reinberg, D. and Roeder, R.G. (1987). Factors involved in specific transcription by mammalian RNA polymerase II. Transcription factor IIS stimulates elongation of RNA chains. J. Biol. Chem. *262*, 3331-3337.

Remacle, J. and Holmberg, S. (1992). A REB1-binding site is required for GCN4-independent ILV1 basal level transcription and can be functionally replaced by an ABF1-binding site. Mol. Cell. Biol. *12*, 5516-5526.

Rhode, P.R., Sweder, K.S., Oegema, K.F., and Campbell, J.L. (1989). The gene encoding ARS-binding factor I is essential for the viability of yeast. Genes Dev. *3*, 1926-1939.

Rigby, R. (1993). Three in one and one in three: it all depends on TBP. Cell *72*, 7-10.

Riggs, A.D., Suzuki, H., Bourgeois, S. (1970). lac repressor-operator interactions. J. Mol. Biol.*48*, 67-83.

Riggs, P.D. (1990). Expression and purification of maltose-binding protein fusions. In Current protocols in molecular biology. F.M. Ausubel, R.E. Brent, and R.P. Perry, eds. (New York, New York: Green Publishing Associates and Wiley Interscience),

Rine, J. and Herskowitz, I. (1987). Four genes responsible for position effect on expression from *HML* and *HMR* in *Saccharomyces cerevisiae*. Genetics *116*, 9-22.

Roeder, R.G. (1991). The complexities of eukaryotic transcription initiation: regulation of preinitiation complex assembly. Trends Gens. *16*, 402-408.

Rose, M.D., Winston, F., and Hieter, P. (1990). Methods in Yeast Genetics (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press).

Ruppert, S., Wang, E., and Tjian, R. (1993). Cloning and expression of human TAFII250: a TBP-associated factor implicated in cell-cylcle regulation. Nature. *362*, 175-179.

Ryder, K., Silver, S., DeLucia, A., Fanning, E., and Tegtmeyer, P. (1986). An altered DNA conformation in origin region I is a determinant for the binding of SV40 large T antigen. Cell *44*, 719-725.

Saltzman, A.G. and Weinmann, R. (1989). Promoter specificity and modulation of RNA polymerase II transcription. FASEB J. *3*, 1723-1733.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular Cloning (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press).

Samuels, M., Fire, A., and Sharp, P.A. (1982). Separation and characterization of factors mediating accurate transcription by RNA polymerase II. J. Biol. Chem. *257*, 14419-14427.

Sanger, F., Nicklen, S., and Coulson, A. (1977). DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA *74*, 5463-5467.

Santangelo, G.M. and Tornow, J. (1990). Efficient transcription of the glycolytic gene *ADH1* and three translational component genes requires the *GCR1* product, which can act through TUF/GRF/RAP binding sites. Mol. Cell. Biol. *10*, 859-862.

Sauer, R., Jordan, S., and Pabo, C. (1990). Lambda repressor: a model system for understanding protein-DNA interactions. Adv. Protein Chem. *40*, 1-61.

Sawadogo, M. (1990). RNA polymerase B (II) and general transcription factors. Annu. Rev. Biochem. *59*, 711-754.

Sawadogo, M. and Roeder, R.B. (1985). Interaction of a gene-specific transcription factor with the Adenovirus major late promoter upstream of the TATA box region. Cell *43*, 165-175.

Schauer, M., Chalepakis, G., Willmann, T. and Beato, M. (1989). Binding of hormone accelaerates the kinetics of glucocorticoid and progesteron receptor binding to DNA.

Schmidt, M.C., Kao, C.C., Pei, R., and Berk, A.J. (1989). Yeast TATA-box transcription factor gene. Poc. Natl. Acad. sci. USA *86*, 7785-7789.

Schmidt, M.C., Zhou, Q., and Berk, A.J. (1989). Spl activates transcription without enhancing DNA-binding activity of the TATA box factor. Mol. Cell. Biol. *9*, 3299-3307.

Schrodinger, E. (1959). Mind and Matter. (Cambridge: Cambridge University Press).

Schroth, G.P., Cook, G.R., Bradbury, E.M., and Gottesfeld, J.M. (1989). Transcription factor IIIA induced bending of the *Xenopus* somatic 5S gene promoter. Nature *340*, 487-488.

Schultz, M.C., Reeder, R., and Hahn, S. (1992). Variants of the TATA-binding protein can distinguih subsets of RNA plymerase I, II, and III promoters. Cell *69*, 697-702.

Schultz, S.C., Shields, G.C., and Steitz, T.A. (1991). Crystal structure of a CAP-DNA complex: the DNA is bent by 90°. Science *253*, 1001-1007.

Schwann, T. (1838). Vorlaufige Milteilung, betreffend Versuch uber die Weingarung und Faulnis. Ann. Phys. *41*, 184-193.

Schwann, T. (1839). Mikroskopische Untersuchungen. (Berlin Sander).

Scott, E.W. (1992). Expression of the *TPI* gene of *Saccharomyces cerevisiae* is controlled by a single complex upstream activating sequence containing binding sites for three *trans*-acting factors: REB1, RAP1, and GCR1. Ph.D. dissertation, University of Florida.

Scott, E.W. and Baker, H.V. (1993).Concerted action of the transcriptional activators REB1, RAP1, and GCR1 in the high-level expression of the glycolytic gene *TPI*. Mol. Cell. Biol. *13*, 543-550.

Scott, E.W., Allison, H.E., and Baker, H.V. (1990). Characterization of *TPI* gene expression in isogeneic wild-type and *gcr1*-deletion mutant strains of *Saccharomyces cerevisiae*. Nucleic Acids Res. *18*, 7099-7107.

Seta, F.D., Ciafre, S-A., Marck, C., Santoro, B., Presutti, C., Sentenac, A., and Bozzoni, I. (1990). The ABF1 factor is the transcriptional activator of the L2 ribosomal protein gene in *Saccharomyces cerevisiae*. Mol. Cell. Biol. *10*, 2437-2441.

Sharp, P.A. (1992). TATA-binding protein is a classless factor. Cell *68*, 819-821.

Sherman, F., Fink, G.R., and Hicks, J.B. (1986). Methods in Yeast Genetics (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press).
Shore, D. and Nasmyth, K. (1987). Purification and cloning of a DNA binding protein from yeast that binds to both silencer and activator elements. Cell *51*, 721-732.

Shore, D., Stillman, D.J., Brand, A.H., and Nasmyth, K. (1987). Identification of silencer binding proteins from yeast: possible roles in SIR control and DNA replication. EMBO J. *6*, 461-467.

Sigler, P.B. (1988). Acid blobs and negative noodles. Nature *333*, 210-212.

Singer, V.L., Wobbe, C.R., and Struhl, K. (1990). A wide variety of DNA sequences can functionally replace a yeast TATA element for transcriptional activation. Genes Dev. *3*, 636-645.

Snyder, U.K., Thompson, J.F., and Landy, A. (1989). Phasing of protein-induced DNA bends in a recombination complex. Nature *341*, 255-257.

Stahl, G.E. (1697). Zymotechnie fundamentalis order allgemeine Grund-erkamtuis der Gahrungs-kunst. (Frankfurt and Liepzig).

Stanway C.A., Chambers, A., Kingsman A.J., and Kingsman, S.M. (1987). Characterization of the transcriptional potency of sub-elements of the UAS of the yeast *PGY* mini-promoter. Nucleic Acids Res. *17*, 9205-9218.

Stanway, C., Mellor, J., Ogden, J.E., Kingsman, A.J., and Kingsman, S.M. (1987). The UAS of the yeast *PGK* gene contains functionally distinct domains. Nucleic Acids Res. *15*, 6855-6873.

Stanway, C.A., Sowden, M.P., Wilson, L.E., Kingsman A.J., and Kingsman, S.M. (1987). Efficient activation of transcription in yeast by the BPV1 E2 protein. Nucleic Acids Res. *17*, 2187-2196.

Stanway, C.A., Chambers, A., Kingsman, A., and Kingsman, S. (1989). Characterization of the transcriptional potency of sub-elements of the UAS of the yeast PGK gene in a PGK mini-promoter [published erratum appears in Nucleic Acid Res. *25*, 10526]. Nucleic Acids Res. *25*,9205-9218.

Stanway, C.A., Gibbs, J.M., Kearsey, S.E., Lopez, M.C., and Baker, H.V. (1994). The yeast co-activator GAL11 positively influences transcription of the phosphoglycerate kinase gene, but only when RAP1 is bound to its upstream activation sequence. Mol. Gen. Genet.*243*, 207-214.

Struhl, K. (1985). Naturally occurring poly(dA-dT) sequences are upstream promoter elements for constitutive transcription in yeast. Poc. Natl. Acad. sci. USA *82*, 8419-8423.

Struhl, K. (1987). Promoters, activator proteins, and the mechanism of transcriptional initiation in yeast. Cell *49*, 295-297.

Suck, D., Lahm, A., and Oefner, C. (1988). Structure refined to 2A of a nicked DNA octanucleotide complex with DNaseI. Nature *332*, 464-468.

Suck, D. and Oefner, C. (1986). Structure of DNaseI at 2.0A resolution suggests a mechanism for binding to and cutting DNA. Nature *321*, 620-625.

Sumimoto, H., Ohkuma, Y., Sinn, E., Kato, H., Shimasaki, S., Horikoshi, M., and Roeder, R.G. (1991). Conserved sequence motifs in the small subunit of human general transcription factor TFIIE. Nature *354*, 401-404.

Sussel, L. and Shore, D. (1991). Separation of transcriptional activation and silencing functions of the *RAP1*-encoded repressor/activator protein 1: isolation of viable mutants affecting both silencing and telomere length. Poc. Natl. Acad. sci. USA *88*, 7749-7753.

Suzuki, M. (1990). The heptad repeat in the largest subunit of RNA polymerase II binds by intercalating into DNA. Nature *344*, 562-565.

Sweder, K.S., Rhode, P.R., and Campbell, J.L. (1988). Purification and characterization of proteins that bind to yeast *ARSs*. J. Biol. Chem. *263*, 17270-17277.

Synder, M., Buchman, A.R., and Davis, R. (1986). Bent DNA at a yeast autonomously replicating sequence. Nature *324*, 87-89.

Sypes, M. and Gilmour, D. (1994). Protein/DNA crosslinking of a TFIID complex reveals novel interactins downstream of the transcription start. Nucleic Acids Res. *22*, 807-814.

Taggart, A., Fisher, T., and Pugh, B. (1992). The TATA-binding protein and associated factors are components of pol III transcription factor TFIIIB. Cell *71*, 1015-1028.

Talanian, R.V., McKnight, C.J., and Kim, P.S. (1990). Sequence-specific DNA binding by a short peptide diamer. Science *249*, 769-771.

Tamura, T., Sumita, K. Fujino, I., Aoyama, A. Horikoshi, M. Hoffmann, A. Roeder, R.G., Muramatsu, M. and Mikoshiba, K. (1991). Striking homology of the 'variable' N-terminal as well as the 'conserved core' domains of the mouse and human TATA-factors (TFIID). Nucleic Acids Res. *19*, 3861-3865.

Tornow, J. and Santangelo, G.M. (1990). Efficient expression of the *Saccharomyces cerevisiae* glycolytic gene *ADH1* is dependent upon a *cis*-acting regulatory element (UAS_{RPG}) found initially in genes encoding ribosomal proteins. Gene *90*, 79-85.

Tornow, J., Xiao Z., Wei, G., and Santangelo, G.M. (1993). GCR1, a transcriptional activator in *Saccharomyces cerevisiae*, complexes with RAP1 and can function without its DNA binding domain. EMBO J. *12*, 2431-2437.

Travers, A.A. (1989). DNA conformation and protein binding. Annu. Rev. Biochem. *58*, 427-452.

Trumbly, R.J. (1992). Glucose repression in the yeast *Saccharomyces cerevisiae*. Molecular Microbiology *6*, 15-21.

Tsang, J.S.H., Henry, Y., Chambers, A., Kingsman, A.J., and Kingsman, S.M. (1990). Phosphorylation influences the binding of the yeast RAP1 protein to the upstream activating sequence of the *PGK* gene. Nucleic Acids Res. *18*, 7331-7337.

Türkel, S. (1994). Transcriptional Regulation of Yeast Retrotransposon Ty2-917. Ph.D. dissertation, University of Maryland.

Uemura, H. and Fraenkel, D.G. (1990). *gcr2*, a new mutation affecting glycolytic gene expression in *Saccharomyces cerevisiae*. Mol. Cell. Biol. *10*, 6389-6396.

Uemura, H. and Jigami, Y. (1992). Role of GCR2 in transcriptional activation of yeast glycolytic genes. Mol. Cell. Biol. *12*, 3834-3842.

VanDyke, M.W. and Sawadogo, M. (1990). DNA-binding and transcriptional properties of human transcription factor TFIID after mild proteolysis. Mol. Cell. Biol. *10*, 3415-3420.

VanHolde, K.E., Lohr, D.E., and Robert, C. (1992). What happens to nucleosomes during transcription? J. Biol. Chem. *267*, 2837-2840.

Van Hoy, M., Leuther, K.K., Kodadek, T., and Johnston, S.A. (1993). The acidic activation domains of the GCN4 and GAL4 proteins are not alpha helical but form beta sheets. Cell. *72*, 587-594.

Vignais, M.L., Huet, J., Buhler, J-M, and Sentenac, A. (1990). Contacts between the factor TUF and RPG sequences. J. Biol. Chem. *265*, 14669-14674.

Vignais, M.L. and Sentenac, A. (1989). Asymmetric DNA bending induced by the yeast multifunctional factor TUF. J. Biol. Chem. *264*, 8463-8466.

Vignais, M.L., Woudt, L.P., Wassenaar, G.M., Mager, W.H., Sentenac, A., and Planta, R.J. (1987). Specific binding of TUF factor to upstream activation sites of yeast ribosomal protein genes. EMBO J. *6*, 1451-1457.

Warner, J.R. (1989). Synthesis of ribosomes in *Saccharomyces cerevisiae*. Micro. Rev. *53*, 256-271.

Weinzierl, R., Ruppert, S., Dynlacht, B., Tanese, N., and Tjian, R., (1993). Cloning and expression of Drosophila TAFII60 and human TAFII70 reveal conserved interactions with other subunits of TFIID. EMBO J. *12*, 5303-5309.

Weiss, M.A., Ellenberger, T., Wobbe, C.R., Lee, J.P., Harrison, S.C., and Struhl, K. (1990). Folding transition in the DNA-binding domain of GCN4 on specific binding to DNA. Nature *347*, 575-578.

White, R. and Jackson, S. (1992). Mechanism of TATA-binding protein recruitment to a TATA-less class III promoter. Cell *71*, 1041-1053.

White, R., Jackson, S., and Rigby, P. (1992). A role for the TATA-box-binding protein component of the transcription factor IID complex as a general RNA polymerase III transcipton factor. Poc. Natl. Acad. sci. USA *89*,1949-1953.

Willett, C., Gelfman, C., and Holland, M. (1993). A complex regulatory element from the yeast gene ENO2 modulates GCR1-dependent transciptional activation. Mol. Cell. Biol. *13*, 2623-2633.

Williams, J.S., Eckdahl, T.T., and Anderson, J.N. (1988). Bent DNA functions as a replication enhancer in *Saccharomyces cerevisiae*. Mol. Cell. Biol. *8*, 2763-2768.

Williams, L.P. and Steffens, H.J. (1978). The History of Science in Western Civilization. Washington D.C.: University Press of America.

Willis, T. (1659). Dr. Willis' Practic of Physick (translated by S. Pordage). (London: Dring, Harper and Leigh).

Wills, C. (1990). Regulation of sugar and ethanol metabolism in *Saccharomyces cerevisiae*. Crit. Rev. Biochem. Mol. Biol. *25*, 245-280.

Wobbe, C.R. and Struhl, K. (1990). Yeast and human TATA-binding proteins have nearly identical DNA sequence requirements for transcription *in vitro*. Mol. Cell. Biol. *10*, 3859-3867.

Wong, J.M., Liu, F., and Bateman, E. (1992). Cloning and expression of the *Acanthamoeba castellanii* gene encoding trancription factor TFIID. Gene *117*, 91-97.

Workman, J.L. and Buchman, A.R. (1993). Multiple functions of nucleosomes and regulatory factors in transcription. Trens. Biochem. Sci. *18*, 90-95.

Wu, H-M. and Crothers, D.M. (1984). The locus of sequence-directed and protein-induced DNA bending. Nature *308*, 509-513.

Yokomori, K., Chen, J., Admon, A., Zhou, S., and Tjian, R. (1993). Molecular cloning and characterization of dTAFII30 alpha and dTAFII30 beta: two small subunits of Drosophila TFIID. Genes Dev. *7*, 2587-2597.

Zhang, X., Zhou, Y., Ebright, Y., and Ebright, R. (1992). Catabolite gene activator protein (CAP) is not an "acidic activating region" transcription activator protein. Negatively charged amino aids of CAP that are solvent-accessible in the CAP-DNA complex play no role in transcription activation at the lac promoter. J. Biol. Chem. *267*, 8136-8139.

Zinkel, S.S. and Crothers, D.M. (1991). Catabolite activator protein-induced DNA bending in transcription initiation. J. Mol. Biol. *219*, 201-215.

BIOGRAPHICAL SKETCH

Michael Andrew Huie was born in 1963 in Munich to Douglas T. Huie and Ingrid A.Huie. He attended college at Columbia University in New York where he recieved an A.B. degree in molecular biology in 1985. He studied medicine at the University of Florida, recieving an M.D.degree in 1993. He also recieved internship training at the University of Florida between 1993 and 1994. He plans to further his training studying dermatology at the University of California, San Francisco begining in 1994.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosphy.

Henry V. Baker, Chair
Associate Professor of Immunology
and Medical Microbiology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosphy.

Richard W. Moyer
Professor of Immunology and
Medical Microbiology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosphy.

Thomas P. Yang
Associate Professor of Biochemistry
and Molecular Biology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosphy.

Daniel Driscoll
Assistant Professor of Immunology
and Medical Microbiology

This dissertation was submitted to the Graduate Faculty of the College of Medicine and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

August 1994

Dean, College of Medicine

Dean, Graduate School

UNIVERSITY OF FLORIDA
3 1262 08554 8161